THREE ESSAYS ON MONETARY POLICY ANALYSIS IN MONGOLIA

Davaajargal Luvsannyam

A thesis submitted for the degree of Doctor of Philosophy of

The Australian National University

June 2012

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of the author's knowledge, it contains no material previously published or written by other person, except where due reference is made in the text.

Signature

Name: Davaajargal Luvsannyam Date: June 20, 2012

ACKNOWLEDGEMENTS

I am indebted to many people for their contributions and support during my study. First of all, I would like to thank the members of my PhD supervisory panel, Professors Ippei Fujiwara, Patrick Fontenay and Dr. Renata Hasanova for their support and helpful feedback. Their guidance helped me throughout the research and writing of this thesis.

I am grateful for comments received at presentations I have made at PhD seminars at Crawford School of Public Policy, the Australian National University.

My appreciation also goes to the staff of the Bank of Mongolia for providing me with the data for this thesis.

I would like to thank all the PhD students in economics at the Crawford School for their encouragement and friendship. I have benefited greatly from their experience.

Financial support from ALA scholarship made it is possible for me focus only on the study and is thus greatly appreciated. I thank Dr. Susanne Holzknecht for copyediting.

Last but not least, I also thank my wife, Munkhtuya Narmandakh, for the support and encouragement she has shown me during the PhD program.

ABSTRACT

This thesis presents three papers on monetary policy analysis for Mongolia. The first paper measures the lagged effect of the monetary transmission mechanism on inflation and output in Mongolia using a sign-restricted structural vector autoregression (SVAR). We find the following results. First, the lag of the monetary transmission mechanism is about 4 to 12 months for Mongolia. Second, monetary policy shocks play a modest role in explaining output and inflation fluctuations. Third, in response to a monetary policy shock, the exchange rate immediately overshoots its long-run equilibrium rate, a finding consistent with Dornbusch's (1976) famous exchange rate overshooting hypothesis. Fourth, the historical decomposition analysis suggests that besides monetary policy shocks, output fluctuations are largely driven by aggregate supply shocks while inflation is largely driven by oil price and money demand (LM) shocks.

The second paper develops an empirical model for inflation in Mongolia using both Bayesian and classical approaches. In particular, we first estimate long-run markup and money demand relationships using cointegration procedures, and then construct a single-equation error correction model of inflation with possible nonlinearity. The main findings of the paper are summarized as follows. First, the main determinant of inflation is the markup, capturing the impact from unit labor costs, petroleum prices, import prices and the exchange rate. Second, money matters for inflation: excess narrow-money supply seems to determine inflation in the long-run if the model uncertainty and nonlinearity are considered, but adjustment to disequilibria is slow. Third, sustained increases in wages together with petroleum price shocks explain the high and volatile inflation in recent years. We also find two inflationary regimes that are characterized by a degree of inflation persistence.

The third paper estimates the reaction function of the Bank of Mongolia using a Bayesian approach. It addresses this issue by estimating the New Keynesian dynamic stochastic general equilibrium (DSGE) model of a small open economy. The main findings of the paper are as follows. First, the monetary policy reaction function is forward looking in terms of the inflation rate. Second, the central bank of Mongolia has implemented a strong anti-inflationary and exchange rate stabilization policy. Third, there is evidence that the Bank of Mongolia does not respond to output.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADF	augmented Dickey-Fuller
ALA	Australian Leadership Awards
ARCH	Autoregressive Conditional Heteroskedasticity
BMA	Bayesian Model Averaging
BOM	Bank of Mongolia
CB	Central Bank
CBB	Central Bank Bills
CPI	Consumer Price Index
DSGE	Dynamic Stochastic General Equilibrium
GDP	Gross Domestic Product
GFC	Global Financial Crisis
GMM	Generalized Method of Moments
LM	Lagrange multiplier
LR	Likelihood ratio
MCI	Monte Carlo integration
MCMC	Markov Chain Monte Carlo
MSE	Mongolian Stock Exchange
MT	Median Target
NSO	National Statistical Office
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
SE	Standard error
SVAR	Structural Vector Autoregression
US	United States
VAR	Vector Autoregression
WTO	World Trade Organization

Chapter 1 INTRODUCTION

It is not surprising that monetary policy issues have been exciting research subjects for both policy makers and academics over the past decade. Interest rate decision making by central banks has an enormous impact on every household and business. Thus it is important to understand how interest rates affect output, employment and inflation. Recent research using a New Keynesian model as a workhorse has attempted to clarify these monetary policy issues. In particular, these studies have identified several channels for monetary policy such as interest rates, exchange rates, inflationary expectations, bank lending, balance sheet effects and wealth effects.

This thesis uses a dynamic stochastic general equilibrium (DSGE) model, a structural vector autoregression (SVAR) model, and time series analysis with special emphasis on the Bayesian approach. Nowadays, a New Keynesian framework that incorporates nominal price or wage rigidity into a DSGE model is an indispensable tool for analysing monetary economics and monetary policy. Many central banks have built their own DGSE models to use for policy analysis and forecasting. For estimation purposes, a classical approach can be used; there are several practical reasons, however, for using the Bayesian approach. First, compared with the classical approach, the Bayesian approach does not suffer from small sample problems. Second, a Bayesian approach comfortably accommodates prior information into the model (Sims 2011). Third, a Bayesian estimation and model comparison are consistent in the case of misspecified models.

Chapter 2 measures the lagged effect of the monetary transmission mechanism on output and inflation using a sign-restricted structural vector autoregression. We find the following results. First, the lag in the monetary transmission mechanism is about 4-12 months for Mongolia. Second, monetary policy shocks play a modest role in explaining output and inflation fluctuations. Third, in response to a monetary policy shock, the exchange rate immediately overshoots its long-run equilibrium rate.

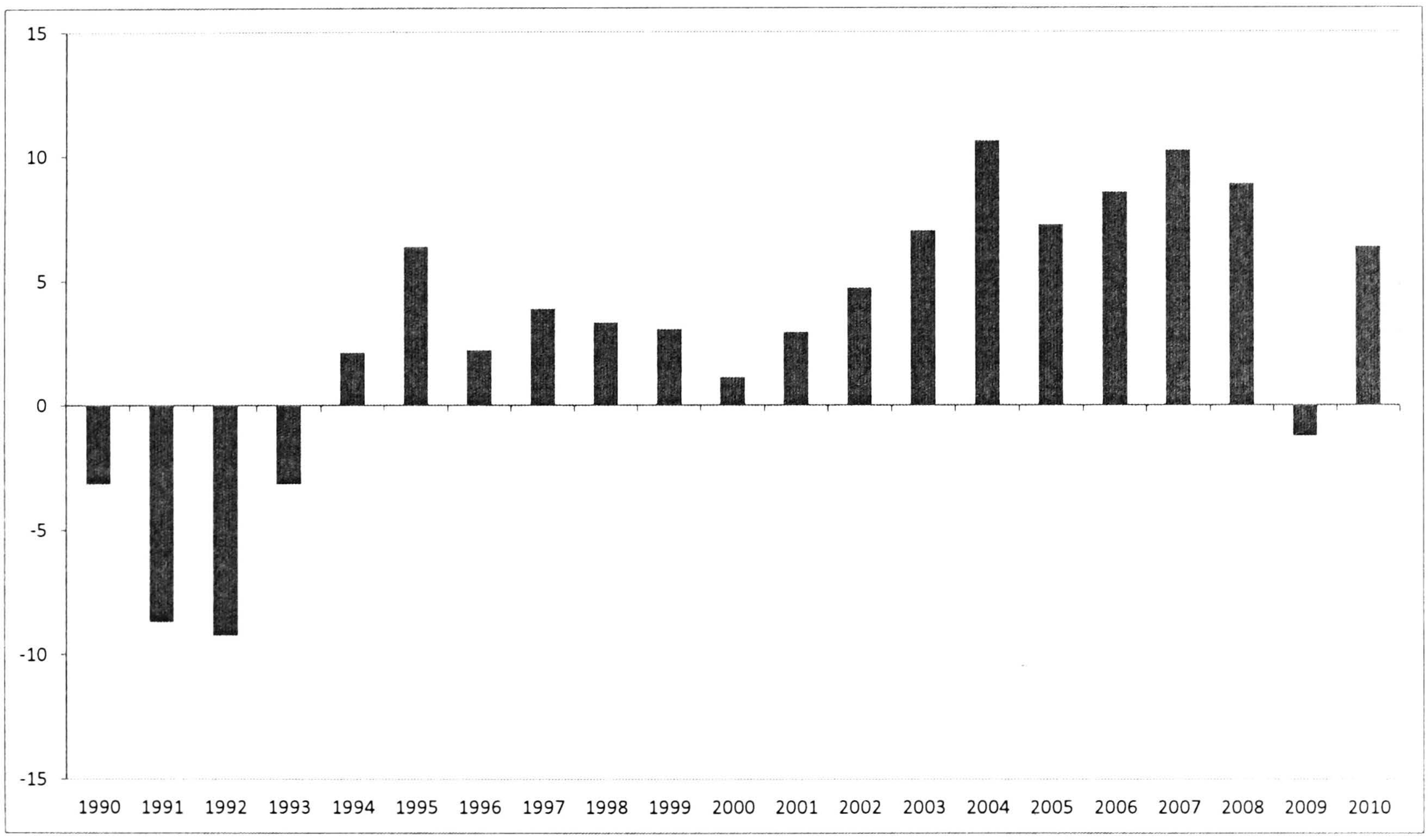
Chapter 3 develops an empirical model for inflation dynamics in Mongolia using both the Bayesian and the Classical approaches. The main findings of the chapter are summarized as follows. First, the main determinant of inflation is the markup, capturing the impact from unit labor costs, petroleum prices, import prices, and the exchange rate. Second, money matters for inflation in both the short and the long-run. Third, sustained increases in wages together with petroleum price shocks explain the high and volatile inflation in recent years. Finally, we also find two inflationary regimes that are characterized by a degree of inflation persistence.

Chapter 4 estimates the reaction function of the Bank of Mongolia (BOM) using a Bayesian approach. We address this issue by estimating a New Keynesian dynamic stochastic general equilibrium (DSGE) model for a small open economy. Our main findings are summarized as follows. First, the monetary policy reaction function is forward-looking in terms of inflation. The expected inflation rule fits the reaction function better than a simple Taylor-type rule. Second, the central bank of Mongolia has implemented strong anti-inflationary and exchange rate stabilization policies. Third, there is evidence that the Bank of Mongolia does not respond significantly to output fluctuations according to the Bayesian posterior odds.

1.1 Economic background of Mongolia

Mongolia is a landlocked country sandwiched between Russia and China, with a population of about 2.7 million and a per capita gross domestic product (GDP) of about US$1,500 in 2007. After the collapse of the Soviet regime in 1990, Mongolia moved towards a market economy. During the first years of the transition period, the country faced major challenges. The breakdown of its economic relationship with the former USSR led to a loss of financial assistance equal to about 60 percent of its GDP. Growth rate in GDP started to fall in 1990 and continued to do so until 1993, only beginning to recover from 1994 onwards. Real GDP reverted to its level prior to the transition in 2001 (see Figure 1.1).

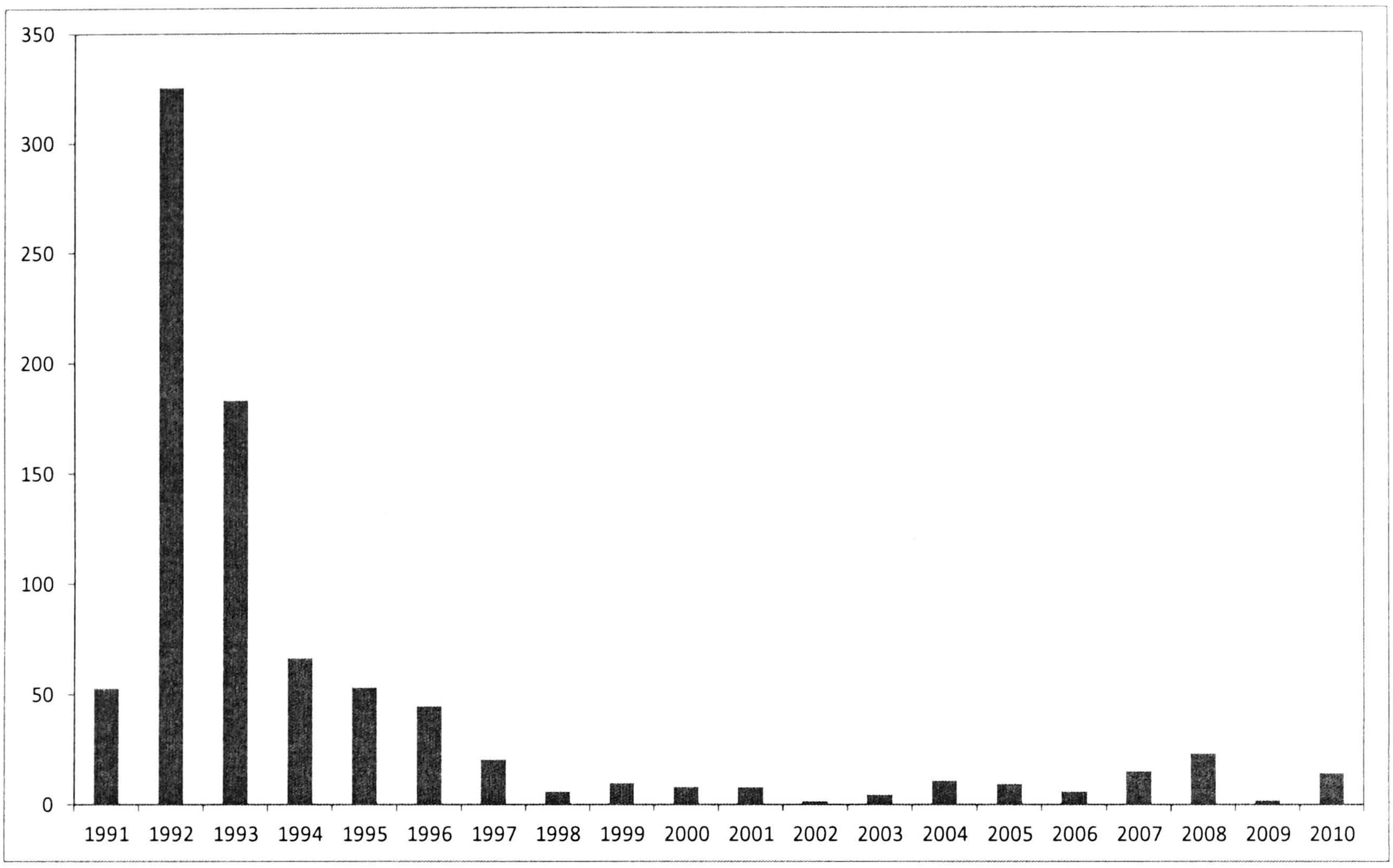
Figure 1.1 Real GDP growth in Mongolia



Source: National Statistical Office of Mongolia

Annual inflation rose sharply from 53 percent in 1991 to a high of 321 percent in 1992. After this crisis, the first sign of economic recovery was observed in 1994. Assistance from international institutions and donors such as the IMF, Japan and the USA played a critical role in the country's recovery (Goyal, 1999). As a result of the tight monetary policy imposed by the Bank of Mongolia, inflation fell markedly to a single digit figure of 6 percent by 1998.

Figure 1.2 Annual inflation in Mongolia



Source: National Statistical Office of Mongolia

Mongolia applied for membership in the World Trade Organization (WTO) in 1991, and joined in 1997; all import tariffs were abolished shortly afterwards. In 2010, Mongolia was trading with 132 countries all over the world and total external trade turnover reached US$6.2 billion, of which exports made up $2.9 billion and imports $3.3 billion. Mongolia's major exports are mineral products: coal, copper, gold and ferrous metals; its main export partners are China, Canada and Russia. The major import commodities are machinery and equipment, vehicles and fuel. The main import partners are Russia, China, South Korea and Japan.

After the end of Communism, Mongolia established a fixed exchange rate system. The Bank of Mongolia devalued the official exchange rate twice, once in June 1991 and again in January 1993. Shortly after that the country's fiscal authorities announced their intention to move to a managed floating exchange rate system.

The economy was hit hard by the global financial crisis (GFC) in 2008. The price of copper, one of Mongolia's main export commodities, dropped by more than 60 percent in late 2008, which in turn caused government revenue to fall and the exchange rate to depreciate sharply. By early 2009, the economy was on the verge of collapse: the central bank was running out of international reserves. Since then, the economy has been recovering rapidly due to the strong growth of investment in the mining sector, an expansion in coal exports, and a surge in commodity prices.

Mongolia's financial system consists largely of commercial banks, which accounted for approximately 96 percent of the country's total assets in 2010. The Mongolian Stock Exchange (MSE) and insurance industry are still in their infancy. The Stock Exchange was established in 1991, implementing the government's plan for the privatization of large state-owned enterprises. When secondary trading began in 1995, the MSE opened its doors to international and domestic investors. During the next decade, 1996-2006, investors were reluctant to trade on the MSE due to the lack of transparency of listed companies. The MSE has seen rapid growth in recent years, however, and in 2010 was the world's best-performing stock market with a growth of 121 percent.

After the collapse of the Soviet regime, Mongolia established a two-tier banking system in 1991. The State Bank of Mongolia became the central bank, the Bank of Mongolia, whose main objective was to ensure price stability. The Bank of Mongolia's functions include management of interest and exchange rates, financial supervision of banks, issuance of banknotes, management of international reserves, and oversight of government borrowing.

The banking system is highly concentrated: the top three banks account for more than 70 percent of market share. Competition is relatively limited in the mostly privately owned banking system, as indicated by relatively high interest rate spreads of around 6 percent in 2011.

Mongolia's banking system is also moderately dollarized, as about one third of bank deposits are denominated in foreign currency. In the early 1990s, dollarization in Mongolia rose sharply as a result of the liberalization of the trading system, the exchange rate and high inflation. Since then, dollarization has fallen to stabilize at around 30 percent.

In 1991 the Bank of Mongolia introduced a reserve requirement and new banking and central banking legislation was passed. Bank-by-bank ceilings and interest rate controls on commercial bank deposits and lending rates were initially introduced, but eventually abandoned with market liberalization. The degree of independence given to the Bank of Mongolia is similar to that of many developed countries' central banks (Slok 2002). The governor of the Bank of Mongolia is appointed by parliament for a six-year period and makes policy decisions independently of the government. There are also some limitations on the government's ability to borrow from the Bank of Mongolia.

From 1995 to 2006 the Bank of Mongolia has used a monetary aggregate targeting framework. Officially, its operating target is reserve money and the intermediate target is M2. In practice, however, the Bank did not follow its monetary targets strictly, but attempted instead to control short term interest rates (see Table 1.1).

Table 1.1.Targeted and actual values for money

Year	M0 growth (%)		M2 growth (%)	
	Target	Actual	Target	Actual
1995		28.7	38.3	32.9
1996		36.5	31.7	25.8
1997		23.1	19.8	32.5
1998		18.7	4.4	-1.7
1999		49.9	10.8	31.6
2000		18.6	11.2	17.6
2001	11.1	8.2	13.6	27.9
2002	21.5	21.9	35.8	42
2003	13.9	14.5	15.2	49.6
2004	20	17	18	20.4
2005	15	19.7	20	34.6
2006	15	35.7	25	34.8

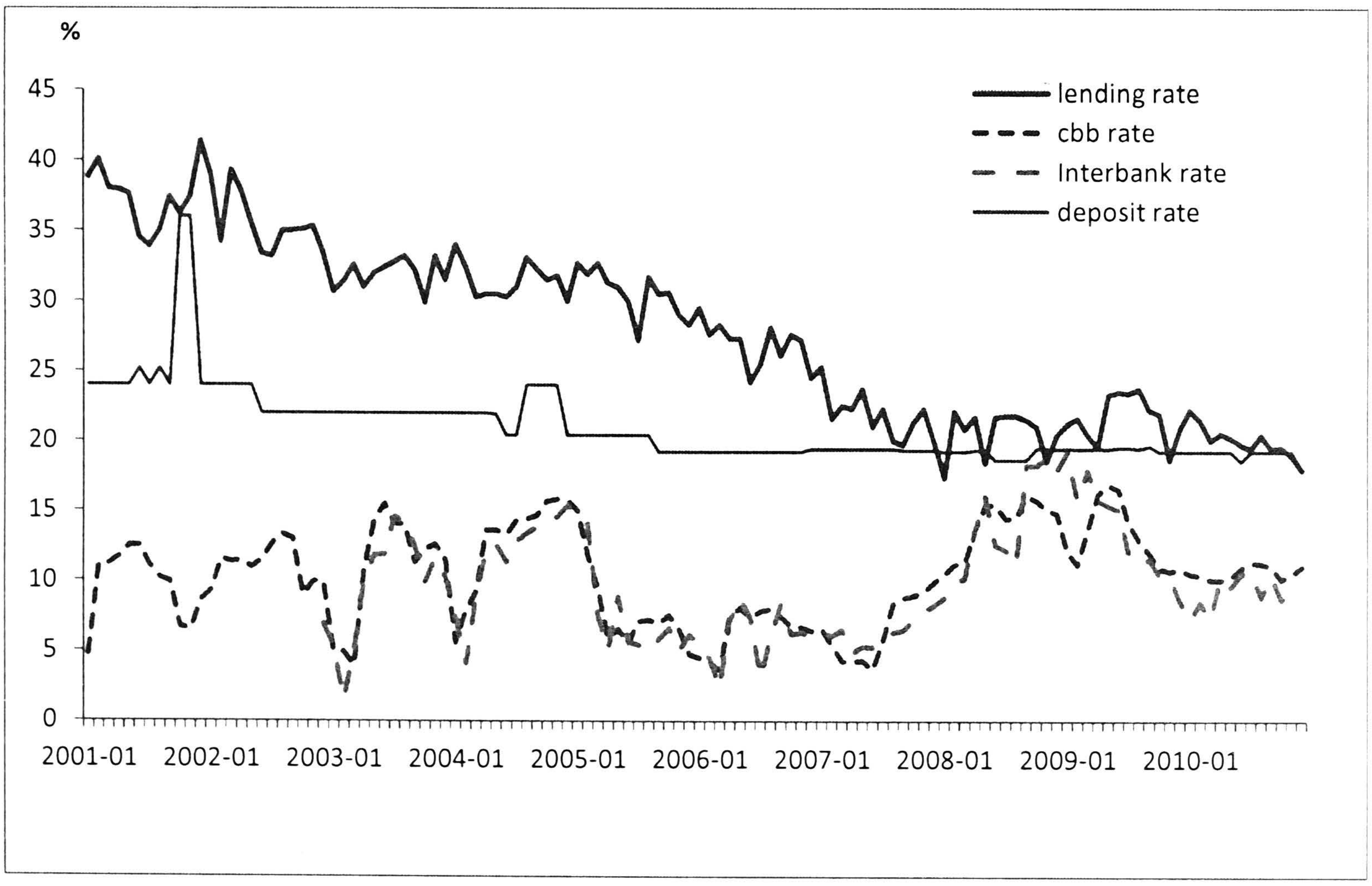
Source: Bank of Mongolia

In 1995, the Bank of Mongolia introduced central bank bills (CBB) as a de facto means of withdrawing liquidity, and since then a regular auction has been held at market determined rates. Central bank bill rates have been very volatile in recent years, fluctuating between 5 and 16 percent. Due to their high volatility and the low transparency on how they are set, it may be difficult for market participants to understand the monetary policy stance of the Bank of Mongolia. In 2007, The Bank changed its monetary aggregate targeting framework to interest rate targeting and announced the seven-day CBB's rate as a policy rate.

Because of the shallowness of the financial markets, the interest rate structure is very different from that of other countries. There is no long-term interest rate such as the Treasury bond rate because the secondary market for government bonds has been underdeveloped for years. High deposit and lending rates have slowly been decreasing in recent years. A lack of integration with international financial markets has led commercial banks and nonbank financial institutions to compete for domestic depositors. Some weak banks also continue to offer unreasonably high interest deposit rates, slowing down the decline of the overall deposit rate. For several of these weak banks, time deposit rates have reached 10-15 percent in real terms. Average weighted lending rates are still more than 18 percent, but this level is not surprising given the high cost of funds, high default risk, and high demand for loans.

Mongolia experienced a strong credit boom with average annual credit growth of 80 percent in 2000-2006. This growth rate is significantly higher than the rates recorded in other transition economies. A credit boom is generally linked to foreign capital inflows, the catching up of loans, economic recovery and a return of confidence in the banking sector. Interest rate transmission between CBB rates and short-term interbank rates has been reasonably strong (Figure 1.3). Short-term interbank rates have closely followed CBB rates. However, CBB rates seem to have a limited impact on bank deposit and lending rates.

Figure 1.3 Interest rate structure in Mongolia



Source: Bank of Mongolia

1.2 Methodology and literature review

Chapter 2 measures the lagged effect of the monetary transmission mechanism on inflation and output in Mongolia using a sign-restricted structural vector autoregression. Recent empirical studies of monetary policy have adopted vector autoregressions (VARs) pioneered by Sims (1980) to estimate the impact of money on the economy was. The development of this approach and the empirical findings that the literature has produced are summarized by Leeper, Sims and Zha (1996). Christiano, Eichenbaum, and Evans (1999) also provide a discussion of the use of VARs to estimate the impact of money, and provide an extensive list of references of work in this area.

In the famous paper that introduced the VAR methodology, Sims (1980) used the Cholesky decomposition. This approach corresponds to a recursive economic structure, or, in other words, the most endogenous variable (in terms of contemporaneous relationships) is ordered last.

In the identification of structural models with contemporaneous restrictions some a priori information is used to impose restrictions on structural VAR models. (Bernanke and Mihov, 1998). A typical restriction compatible with virtually all macroeconomic models is that, in the long-run, demand shocks have zero impact on output. Blanchard and Quah (1989) showed how this type of restriction can be used to identify VARs.

An OLS estimation of a VAR with a number of variables greater than six is typically inefficient due to the small degree of freedom. In order to solve this over-parameterisation problem, a Bayesian VAR was proposed by Litterman (1980) and Doan, Litterman and Sims (1984).

According to Canova (2007), there are two approaches to estimate a structural Bayesian VAR. A naive one, employed by Canova (1991) and Gordon and Leeper (1994), is to use Normal-Wishart priors for reduced form parameters, and draws for the structural parameters are made conditional on the identification restrictions. When the contemporaneous matrix is over-identified, it is better to work with the Sims-Zha model (1998).

Faust (1998), Uhlig (2005), and Canova and De Nicolo (2002) used sign restrictions of the impulse response to identify structural shocks. Compared with recursive (Cholesky decomposition) or short-run identification (contemporaneous), sign-restricted SVARs do not impose zero-type dubious restrictions on the contemporaneous matrix. Instead, this method achieves identification by explicitly restricting the sign of the structural impulse responses using economic theory or a dynamic stochastic general equilibrium approach.

There are a number of approaches to isolate structural shocks using sign-restricted SVARs: the sequential orthogonalization method using the Gram-Schmidt process (Mountford and Uhlig, 2009); the Givens rotation matrix approach (Canova and De Nicolo, 2002); and the Householder (QR) transformation approach (Rubio-Ramirez, Waggoner and Zha, 2005). Fry and Pagan (2011) recommended that practitioners use the Givens rotation matrix or Householder (QR) transformation approach, even when dealing with a single shock.

Although VAR models have acquired an important place in applied macroeconomic research methods, VAR identifications of monetary policy shocks have been criticized on several grounds. First, Cooley and LeRoy (1985) criticize the recursive identification approach because contemporaneous recursive structures are hard to defend from general equilibrium model perspectives. Second, Faust and Leeper (1997) argue that the long-run restrictions approach inaccurately estimates long-run effects of shocks in finite samples, which in turn transfers this imprecision to other parameters of the model.

Third, Fry and Pagan (2011) reviewed critically the sign-restricted VAR literature. In particular, they warned about the multiple shocks and multiple models problem.

Chapter 3 develops an empirical model for inflation in Mongolia. In particular, we first estimate long-run markup and money demand relationships using cointegration procedures and then construct a single-equation error correction model. Money demand and price markup are estimated by classical and Bayesian cointegration approaches, suggested by Johansen (1988), Villani (2005) and Warne (2006). There are a number of Bayesian approaches to cointegration in the literature, including: Kleibergen and van Dijk (1994); Bauwens and Lubrano (1996); Geweke (1996); Bauwens and Giot (1998); Kleibergen and Paap (2002); Strachen (2003); Strachen and Inder (2004); Villani (2005); and Warne (2006). Villani (2005) suggest a prior for cointegration space rather than a prior for the exactly identified cointegration relations. The paper by Warne (2006) extends the Villani cointegration procedure in two important dimensions: first, it allows for proper prior distribution of the short-run parameters on lagged endogenous variables; second, an analytical expression of the posterior mode is derived.

Bayesian model averaging has become an important tool in empirical research with a large number of regressors and a relatively limited number of observations. Here we investigate short run inflation dynamics using general-to-specific modelling (Campos, Ericsson and Hendry, 2005) and the Bayesian model averaging approach developed by Fernandaz, Ley and Steel (2001), who propose a "benchmark" prior distribution that works for the general condition that includes substantive prior information in the analysis.

We model the possible nonlinearity of the inflation persistence using the Markov switching model. Use of this model has proliferated since Hamilton's (1989) seminal paper on business cycle dating. A nice overview of Markov-switching models can be found in Hamilton (1994). The monographs of Kim and Nelson (1999) and Frühwirth-Schnatter (2006) provide detailed techniques of the Bayesian approach for Markov switching models.

Chapter 4 estimates the reaction function of the Bank of Mongolia using a Bayesian approach. Lane (1999) provides an excellent survey of earlier work on optimizing open economy models with nominal rigidities that focus on the transmission of monetary policy shocks. The main contribution in that area is Obstfeld and Rogoff (1995, 1996) who develop a two country model where monopolistically competitive firms set prices one period in advance.

Recent papers such as Obstfeld and Rogoff (2002), Benigno and Benigno (2003), Sutherland (2003), Devereux and Engel (2003), and Corsetti and Pesenti (2005) have focused on the implication of two country sticky price open economy models for the design of optimal monetary policy using a welfare approach More recent frameworks have adopted the staggered price setting structure derived from Calvo (1983). The assumption of staggered price and wage setting introduces more realistic dynamics than that of price setting one period in advance. One unsatisfactory feature of price setting one period in advance is its implication that only unanticipated monetary policies have any effect on real output. Gali and Monacelli (2005), Clarida, Gali and Gertler (2001), Kollmann (2002) and Monacelli (2005) develop small open economy models using the Calvo price setting.

Finally, Clarida, Gali and Gertler (2002), Pappa (2004) and Benigno and Benigno (2006) analyse the alternative monetary policy arrangement in a two country framework with a Calvo staggered price setting and a focus on the gains from cooperation.

The model for this paper is based on a small open economy model, as developed by Gali and Monacelli (2005). However, for estimation purposes it modifies this model after the manner of Lubik and Schorfheide (2007), because a stochastic singularity problem arises if the number of shock terms in the model is less than the number of observed variables.

Chapter 5 concludes the main findings and draws some policy implications.

Chapter 2 [UN]IMPORTANCE OF MONETARY POLICY SHOCKS IN MONGOLIA

2.1 Introduction

Since the 1990s, central banks have changed their monetary policy framework, setting the maintenance of low rates of inflation as their primary objective and announcing inflation targets explicitly, or implicitly, to the general public. In this regard, central banks have tried to determine in detail the issue of the lagged effect of monetary policy on the economy. Monetary policy is transmitted through many channels and after a certain period affects output and inflation. Many channels have been identified, including interest rates, the exchange rate, inflationary expectations, bank lending, balance sheet effects, and wealth effects. However, there is little agreement over their precise workings or relative importance (Mishkin, 1995). For developed countries, the lag period of the monetary transmission mechanism is about 12 to 24 months. In contrast, for a developing country such as Mongolia, the lag period should be much less on account of the relatively underdeveloped financial markets and small, open economy. The reaction of prices and output is faster in small open economies because the exchange rate responds faster to changes in monetary policy.

This paper attempts to measure the lagged effect of the monetary transmission mechanism on inflation and output in Mongolia using a sign-restricted structural vector autoregression (SVAR) developed by Faust (1998), Uhlig (2005) and Canova and De Nicolo (2002). Compared with the traditional recursive method, a sign-restricted SVAR has several advantages. First, it does not impose dubious zero-type restrictions on the contemporaneous matrix. Second, all constraints employed are explicitly stated. Third, this approach avoids the puzzle problems sometimes found in the literature (for example, see Sims, 1992). These puzzles arise if the reactions of the other variables do not appear as they should – if, for instance, a contractionary monetary policy shock is followed by a rise in price level. Fourth, the results are indifferent to a reordering of variables or choice of a different Cholesky decomposition.

By applying a sign-restricted SVAR to a data set for the Mongolian economy from 1996:12 to 2009:12, we find the following results. First, the lag of the monetary transmission mechanism is about 4 to 12 months; contractionary monetary policy which increases the interest rate by 1 percentage point reduces real GDP by about 1.5 percent in four months, while the price level measured by the CPI reacts relatively slowly, with prices dropping by 1.1 percent over a year. Second, monetary policy shocks play a modest role in explaining output and inflation fluctuations. Third, following a monetary policy shock, the exchange rate immediately overshoots its long-run equilibrium rate, a finding that is consistent with Dornbusch's (1976) famous exchange rate overshooting hypothesis. However, according to variance decomposition analysis it is difficult to explain exchange rate volatility by monetary policy shocks. Fourth, the historical decomposition analysis suggests that besides monetary policy shocks, output fluctuations are largely driven by aggregate supply shocks and inflation is largely driven by oil price and LM shocks.

The paper consists of five sections. Section 2.2 reviews and explains the structural vector autoregression model. Empirical results and robustness analysis are shown in sections 2.3 and 2.4, and the last section concludes the paper.

2.2 Structural Vector Autoregression

In the famous articles of Lucas (1976) and Sims (1980), the authors argued that the Cowles Commission or structural models are theoretically incorrect. In particular, they criticized that structural models do not represent data and economic theory and, so are ineffective for the practical purposes of forecasting and policy evaluation. In response to these critiques, the LSE (London School of Economics), Vector Autoregression (Sims, 1980), and Real Business Cycle approaches were proposed.

Recent empirical studies of monetary policy have adopted vector autoregressions (VAR). The use of VARs to estimate the impact of money on the economy was pioneered by Sims (1980). The development of the approach and the empirical findings that this literature has produced are summarized by Leeper, Sims and Zha (1996). Christiano, Eichenbaum and Evans (1999) also provide a discussion of the use of VARs to estimate the impact of money, and provide an extensive list of references of work in this area[1].

A structural vector autoregression with k variables is given as follows:

$$A\begin{pmatrix} Y_t \\ M_t \end{pmatrix} = C(L)\begin{pmatrix} Y_{t-1} \\ M_{t-1} \end{pmatrix} + B\begin{pmatrix} v_t^Y \\ v_t^M \end{pmatrix} \tag{2.1}$$

where:

Y is the vector of non-policy macroeconomic variables (e.g. output and prices);

M is the vector controlled by the monetary policymaker (e.g. interest rates and monetary aggregates containing information on monetary policy actions).

Matrix A describes the contemporaneous relations among the variables and $C(L)$ is a matrix of finite-order lag polynomial. $v \equiv \begin{pmatrix} v^Y \\ v^M \end{pmatrix}$ is a vector of structural disturbances to

[1] A good overview of structural VAR is provided by Amisano and Giannini (1997) and Favero (2001).

the non-policy and policy variables (with unit variance and independent of each other); the non-zero off-diagonal elements of B allow some shocks to affect directly more than one endogenous variable in the system.

The structural model (2.1) is not directly observable, however, a VAR can be estimated as the reduced form of the underlying structural model:

$$\begin{pmatrix} Y_t \\ M_t \end{pmatrix} = A^{-1}C(L)\begin{pmatrix} Y_{t-1} \\ M_{t-1} \end{pmatrix} + \begin{pmatrix} u_t^Y \\ u_t^M \end{pmatrix} \tag{2.2}$$

where u denotes the VAR residual vector, which is independently distributed through time with a full contemporaneous variance-covariance matrix Σ.

The relationship between the VAR residuals, ***u***, and structural disturbances, *v*, is as follows:

$$A\begin{pmatrix} u_t^Y \\ u_t^M \end{pmatrix} = B\begin{pmatrix} v_t^Y \\ v_t^M \end{pmatrix} \tag{2.3}$$

If undoing the partitioning, (2.3) is as follows:

$$u_t = A^{-1}Bv_t \tag{2.4}$$

From which we can derive the relationship between the variance-covariance matrix of u_t and the variance-covariance matrix of v_t as follows:

$$E(u_t u_t^{'}) = A^{-1}BE(v_t v_t^{'})B^{'}A^{-1} \tag{2.5}$$

Substituting population moments with sample moments we have:

$$\hat{\Sigma} = \hat{A}^{-1}\hat{B}I\hat{B}^{'}\hat{A}^{-1} \tag{2.6}$$

The $\hat{\Sigma}$ matrix contains $\frac{n(n+1)}{2}$ different elements and this is the maximum number of identifiable parameters in matrices A and B. In practice, identification requires the imposition of some restrictions on the parameters of the A and B matrices, and depending on how the restriction is imposed, identification can be classified as follows:

1. Cholesky decomposition;
2. Structural model with contemporaneous restrictions;
3. Structural model with long-run restrictions;
4. Sign restrictions of the impulse responses.

In the famous paper that introduced the VAR methodology, Sims (1980) used the following identification strategy, based on the Cholesky decomposition of matrices:

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn-1} & 1 \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix} \tag{2.7}$$

This approach corresponds to a recursive economic structure, or, in other words, the most endogenous variable (in terms of contemporaneous relationships) is ordered last. In the identification of structural models with contemporaneous restrictions some a priori information is used to impose restrictions on the elements of matrices A and B (Bernanke and Mihov, 1998).

A typical restriction compatible with virtually all macroeconomic models is that, in the long-run, demand shocks have zero impact on output. Blanchard and Quah (1989) showed how this type of restriction can be used to identify VARs. The structural model with long-run restrictions is specified by positing that A is equal to the identity matrix and by not imposing any zero restriction on the B matrix:

$$y_t = \sum_{i=1}^{p} A_i y_{t-i} + Bv_t \tag{2.9}$$

From this it is possible to derive the matrix that describes the long-run effect of the structural shocks on the variable of interest as follows:

$$\left(I - \sum_{i=1}^{p} A_i \right)^{-1} Bv_t = -\Pi^{-1} Bv_t \tag{2.10}$$

One problem with the VAR is that there are a large number of free parameters to be estimated. In fact, the number of parameters to be estimated in a VAR with order of p is equal to $n^2 p + n(n+1)/2$. Thus, an OLS estimation of a VAR with a number of variables greater than 6 is typically inefficient due to the small degree of freedom. In order to solve this over-parameterisation problem, a Bayesian VAR was proposed by Litterman (1980) and Doan et al (1984).

There are two approaches to estimating a structural Bayesian VAR according to Canova (2007). A naive one, employed by Canova (1991) and Gordon and Leeper (1994), is to use Normal-Wishart priors for reduced-form parameters, then draws for the structural parameters are made conditional on the identification restrictions. This approach is appropriate if the A matrix is just identified. When the A matrix is over-identified, it is better to work with the Sims-Zha model (1998). The Sims-Zha approach has three important features. First, structural BVAR analysis under informative prior is feasible for large systems. Second, they introduced dummy observation priors of unit root and cointegration into the structural BVAR model. Third, the posterior of the A matrix cannot be computed analytically. To simulate a posterior we need to use a Monte Carlo integration method such as importance sampling or a Metropolis-Hasting algorithm and the restricted Gibbs sampler of Waggoner-Zha (2003).

Recently Faust (1998), Uhlig (2005), and Canova and De Nicolo (2002) used sign restrictions of the impulse response to identify structural shocks. Compared with recursive (Cholesky decomposition) or short-run identification (contemporaneous), sign-restricted SVAR do not impose dubious zero-type restrictions on a contemporaneous matrix. Instead, this method achieves identification by explicitly restricting the sign of structural impulse responses using economic theory or a dynamic stochastic general equilibrium approach.

Although all of the authors used sign-restricted SVAR, Faust (1998), Uhlig (2005), and Canova and De Nicolo (2002) have different justifications. Faust (1998) and Canova and De Nicolo (2002) challenged the robustness of the consensus conclusion based on recursive identification that monetary shocks explain a small share of output fluctuation. In particular Faust (1998) maximizes the forecast-error-variance share of GDP with respect to unit sphere vector and sign restriction on the impulse response of variables to monetary policy shocks. Canova and De Nicolo (2002) identified monetary disturbances by imposing sign restrictions on the cross correlations of variables in response to shocks.

Uhlig (2005) focused on the effects of monetary policy on output and proposed two different but related approaches based on a Bayesian method: pure-sign restriction and penalty function. The pure-sign restrictions algorithm is implemented in the following steps. First, taking a joint draw from the posterior for Normal-Wishart posterior for the VAR parameters ($A^{-1}C(L), \Sigma$) and uniform distribution over the unit sphere (α). Second, calculating the impulse response responses $r_{k,j}$ at horizon $k = 0,\ldots,K$ for the variable j using the impulse vector ($a = \tilde{A}\alpha$) which is equal to the lower triangular Cholesky factor of Σ ($\tilde{A}$) times n dimensional unit length vector α. Third, if all impulse responses for the variables satisfy the sign restrictions, the draw is kept, otherwise it is discarded. Steps 1-3 are then repeated many times and finally, the 16^{th}, 50^{th} (median) and 84^{th} percentiles are calculated based on the draws kept.

The penalty function approach identifies structural shocks by minimizing some penalty function that penalizes positive responses in linear proportion and rewards negative responses in linear proportion. In particular, Uhlig (2005) defined the penalty function as follows:

$$f(x) = \begin{cases} x & if\ x \leq 0, \\ 100x & if\ x \geq 0 \end{cases}$$

Initially the sign-restricted SVAR approach focuses on isolating single shocks. However, this method can easily be generalized to the identification of multiple, uncorrelated structural shocks. There are a number of approaches to isolating structural shocks: a sequential orthogonalization method using the Gram-Schmidt process (Mountford and Uhlig, 2009); the Givens rotation matrix approach (Canova and De Nicolo, 2002); and the Householder (QR) transformation approach (Rubio-Ramirez et al., 2005). Fry and Pagan (2011) recommended that practitioners use the Givens rotation matrix or Householder (QR) transformation approach, even when dealing with a single shock.

Although VAR models have acquired an important place in applied macroeconomic research methods, VAR identifications of monetary policy shocks have been criticized on several grounds. First, Cooley and LeRoy (1985) criticize the recursive identification approach because contemporaneous recursive structures are hard to defend from general equilibrium model perspectives. Second, Faust and Leeper (1997) argue that the long run restrictions approach inaccurately estimates long-run effects of shocks in finite samples which in turn transfers this imprecision to other parameters of the model. Third, Fry and Pagan (2011) critically reviewed the sign-restricted VAR literature. In particular, they warned about the multiple shocks and multiple models problem: the multiple shocks problem arises if there is a failure to impose enough sign restriction to differentiate between shocks, while the multiple models problem arises if there is a difference between the median responses and the median target (MT) proposed as diagnostic tools because the median responses may come from different models. They also argued that researchers should be clear about whether shocks are transitory or permanent since the appropriate summative model depends on the nature of shocks. However, this may not be a problem for Bayesian VAR, which is robust to the presence of non-stationarity and cointegration (see Sims and Uhlig, 1991).

2.3 Empirical results

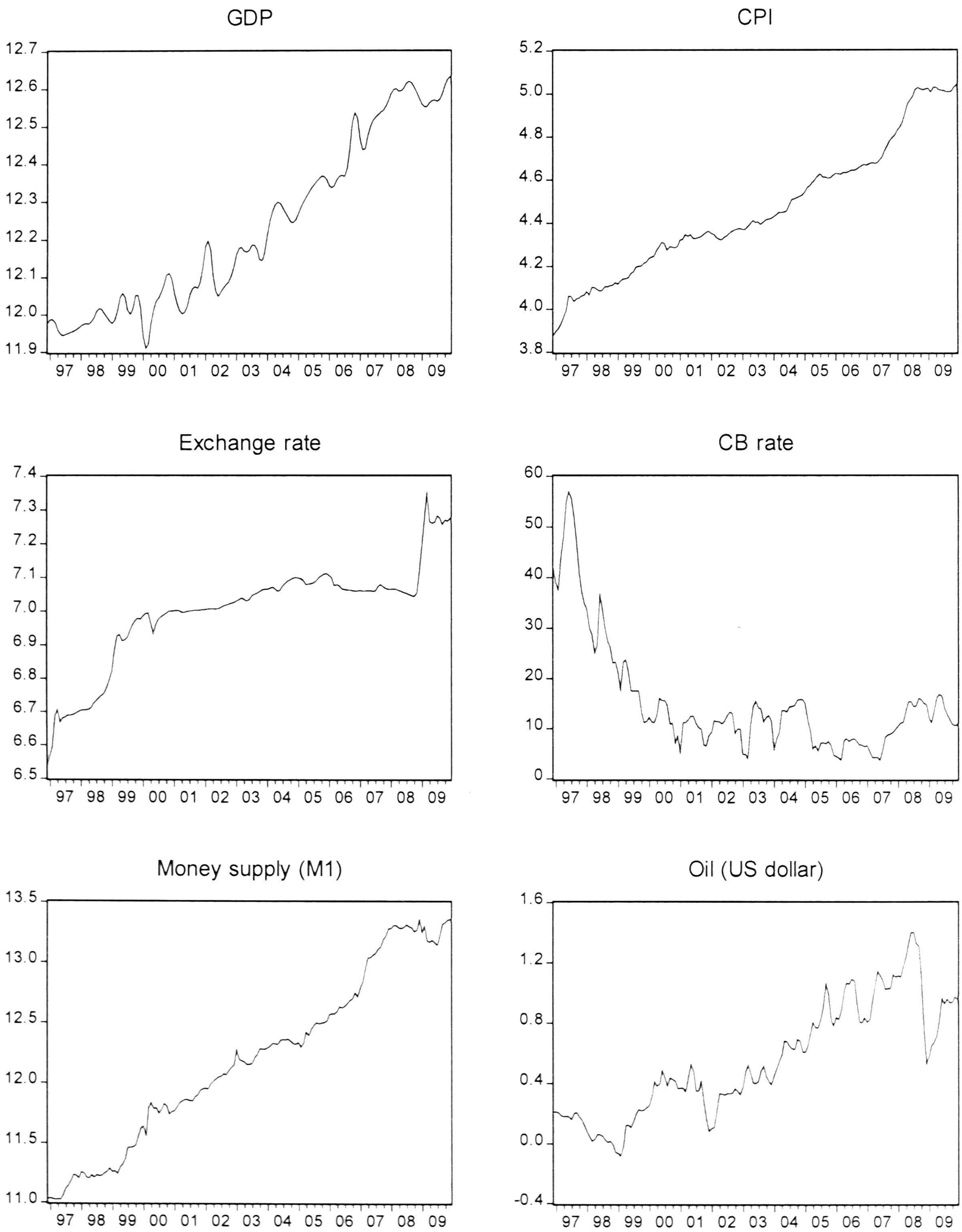
Data

The lagged effects of the monetary transmission mechanism on inflation and output were estimated using monthly data from 1996:12 to 2009:12. The vector autoregression model used to estimate the effects of monetary policy included six variables: oil price, output, price, exchange rate, interest rate, and narrow money. The consumer price index and real GDP (at a constant price of 2005) are taken from the bulletin of the National Statistical Office of Mongolia. Since the frequency of the real GDP data is quarterly, we estimate monthly real GDP using an interpolation method (Doan 2010). Data on the exchange rate (togrog[2] against US dollar), money supply (M1), and central bank bill rate are taken from the monthly bulletin of the Bank of Mongolia. Oil price is the spot oil price in US dollars per barrel taken from the Federal Reserve Board of St Louis website, http://research.stlouisfed.org/. Our measures of output, price, exchange rate, interest rate and narrow money are real GDP (LGDP), consumer price index (LCPI), midpoint togrog rate against the US dollar of the BOM (LEXR), central bank bill rate

[2] The togrog is the official currency of Mongolia

(CBB), and money M1 (LM1) respectively (Figure 2.1). All variables except interest rate are in logarithms and real GDP, CPI, and narrow money are seasonally adjusted using the Census X12 approach.

Figure 2.1 Data



Source: National Statistical Office, Bank of Mongolia and Federal Reserve Board of St Louis

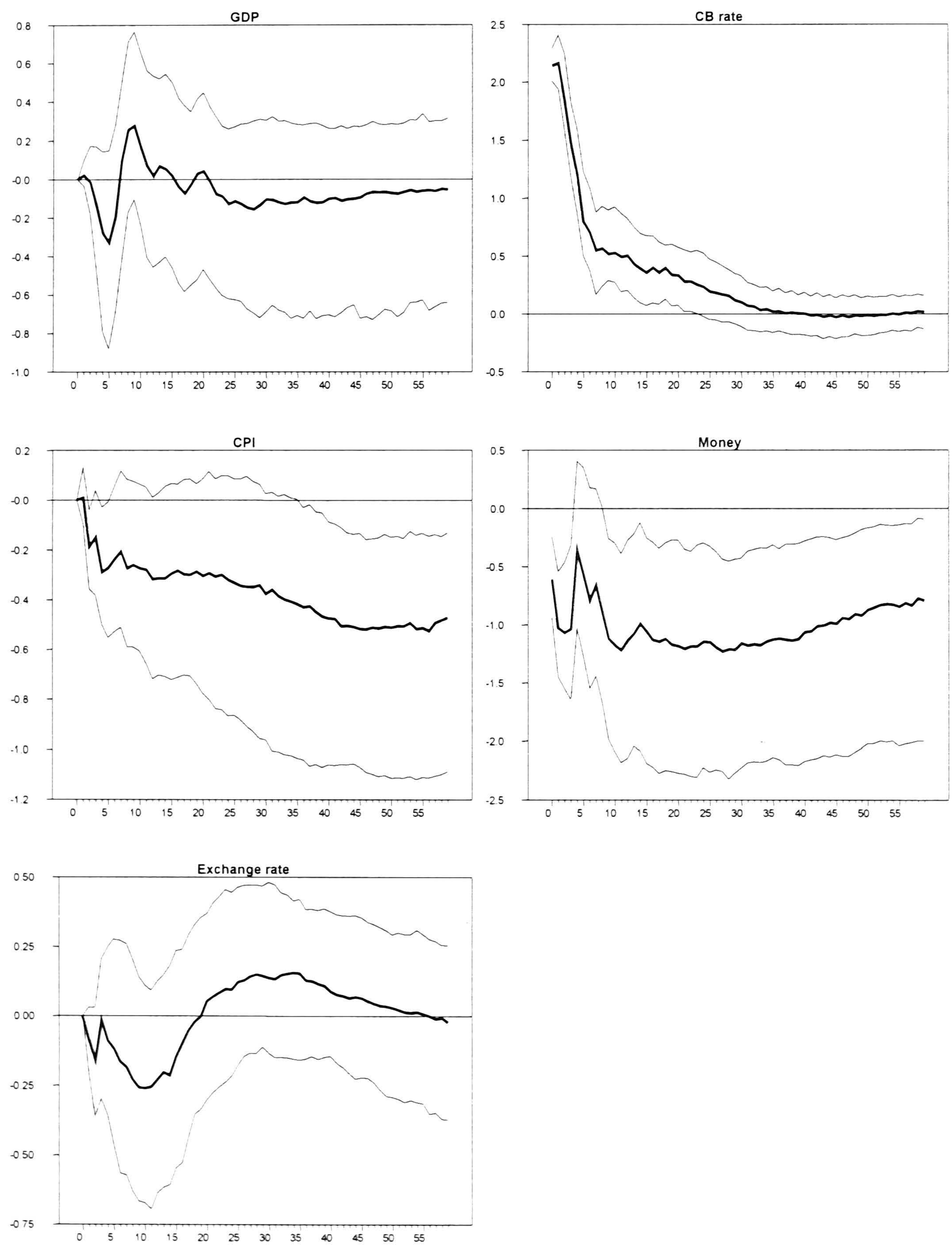
Traditional Cholesky decomposition

As reduced-form residuals are typically correlated, the Cholesky decomposition isolates underlying structural shocks (uncorrelated) by recursive orthogonalization. Since this identification depends on the structure of matrix A, variables are ordered according to their level of endogeneity (the most endogenous variable is ordered last): oil price ⇨ output ⇨ price ⇨ exchange rate ⇨ interest rate ⇨ narrow money. According to the Akaike information criterion, lag length is chosen at p=6 (see Table 2.2a in the appendix). Since the effective sample period covers a time series from June 1997 to December 2009, the number of observations is 151 after the lag adjustment. A constant or time trend is not included in the VAR estimation.

The analytical derivates method of confidence interval is based upon a linearization that becomes increasingly inaccurate as the number of steps grows. Here we use Monte Carlo integration (MCI) methods (Sims and Zha, 1999) to estimate a confidence error band instead. In particular, we apply this approach for the structural VAR with a Normal-Wishart prior and identify structural shocks using the Cholesky factorization. The median estimates of responses are shown as solid (black) lines while thin (blue) lines display a 68 percent error band estimated by the Monte Carlo integration method.

Early papers that identified the effects of monetary policy shocks in small open economies using contemporaneous restriction include Cushman and Zha (1997) and Dungey and Pagan (2000). Starr (2005) studied whether monetary policy variables affect output and price in transition economies using Cholesky decomposition and found mixed evidence. Figure 2.2 shows the effects of monetary policy shocks on macro and financial variables. Although signs of the impulse response function seem to be acceptable with conventional wisdom that monetary contractions should raise interest rates, lower prices, reduce real output, and lead to an appreciation of the exchange rate, the responses of the variables are statistically insignificant except for narrow money (M1). The Cholesky decomposition method fixes the initial responses of real output, prices, and the exchange rate to zero.

Figure 2.2 The impulses responses to a monetary policy shock (Cholesky decomposition)



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

Sign-restricted SVAR

Table 2.1 shows the sign restriction imposed on the impulse responses of the five identified structural shocks. Monetary policy shocks are related to three possible sources: exogenous changes to the preferences of central bankers; exogenous changes in policy induced by changes in private agent's inflationary expectations; and measurement error in the real-time data. A contractionary monetary policy shock is identified by restricting the impulse responses of output, price, exchange rate and money supply to non-positive, and the impulse responses for central bank rate to non-negative. A positive oil price shock is defined as a shock where the dollar price of oil rises. An aggregate supply shock is attributed to the exogenous variations in markup, productivity and other supply-side factors. An aggregate demand shock reflects exogenous impacts of wage-push inflation, fiscal policy and other demand-side factors. A positive aggregate supply shock is identified as a shock where output increases and the price level decreases, while positive aggregate demand shocks are shocks where output and the price level increase. Finally, an LM shock is defined as a shock where money demand increases. In the dominant framework-New Keynesian model, monetary aggregates do not affect IS and Phillips curve equations and the policy reaction equation does not include money. However, there has been interest in analyzing the role of money in the business cycle in recent years. Favara and Giordani (2009) and Canova and Menz (2011) have argued empirically and theoretically that shocks to monetary aggregates have substantial and persistent effects on output, prices and interest rates.

Table 2.1 Identifying sign restriction

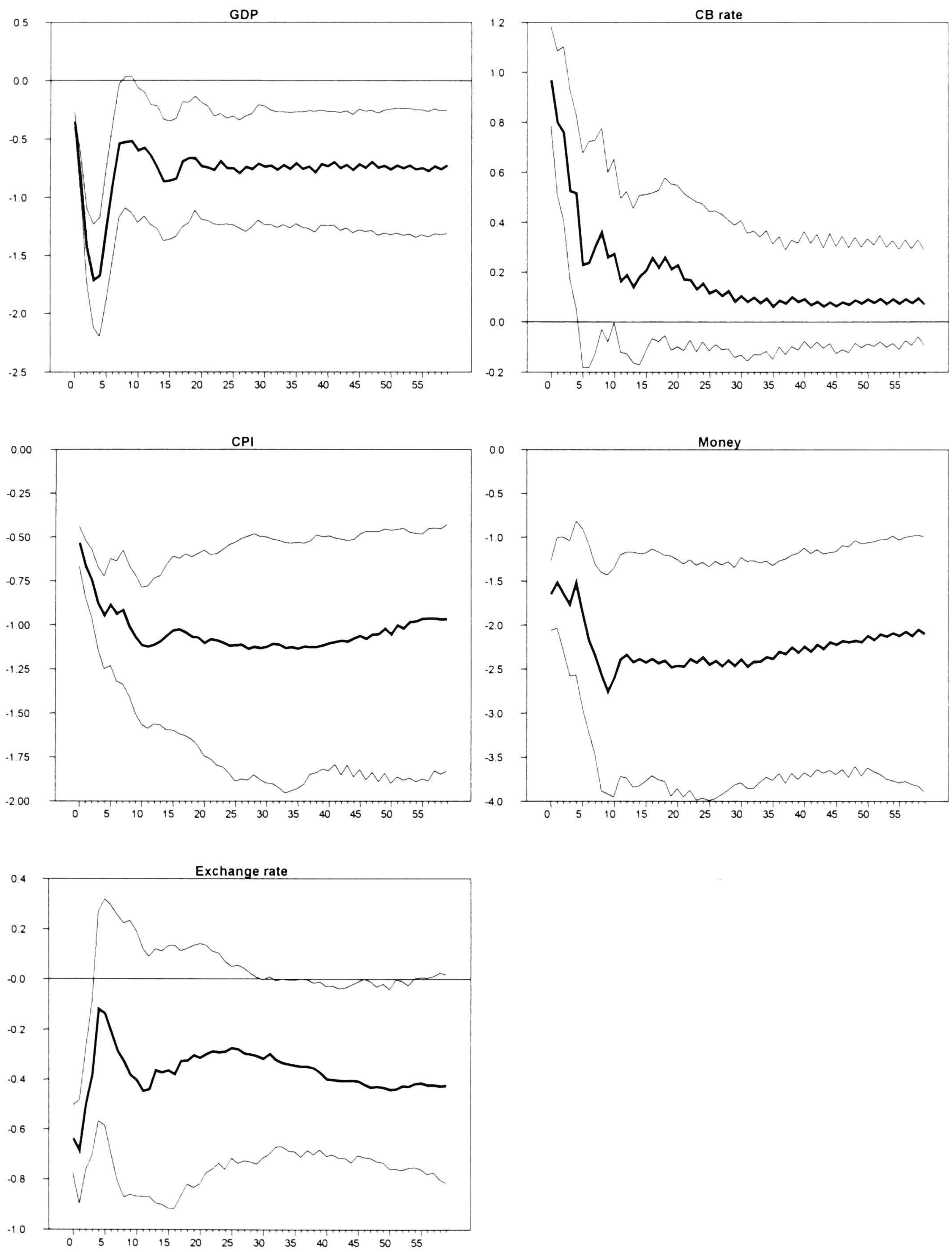
	GDP	CPI	Exchange rate	Interest rate	M1	Oil price
Monetary policy shock	–	–	–	+	–	
Oil price shock						+
Supply shock	+	–				
Demand shock	+	+				
LM shock					+	

Notes: A '+' means that the impulse response of the variable is restricted to be positive for the impact period of the shock. Similarly, a '–' indicates a negative response. A blank entry indicates that no restrictions have been imposed. The sign restriction horizon K=0.

The penalty function approach is used to estimate the impulse responses of the variables. The impulse responses to a monetary policy shock are shown in Figure 2.3. The central bank rate reacts immediately, increasing by 1 percentage point, then reversing course within a year. A rise of 100 basis points in the interest rate reduces real GDP by about 0.6 percent, with the peak effect occurring after four months. This short lag for the effect of a tightening of monetary policy on GDP may seem implausible, but it can be explained by the effect of the overshooting of the exchange rate on a very open economy since Mongolia's import to GDP ratio is about 60 percent. Also this result is consistent with some studies for the Commonwealth Independent Countries (CIS). For example, Starr (2005) found that innovation of the interest rate is associated with a significant decrease in output in one quarter. Compared with output, the CPI goes down immediately then reacts relatively slowly, with prices dropping by 1.1 percent within a year. Money supply (M1) falls immediately, with a maximum lagged effect of 2.5 percent.

Following a contractionary monetary policy shock, the exchange rate appreciates immediately then, within one to two months, gradually depreciates back to equilibrium. The results are therefore consistent with Dornbusch's (1976) well-known exchange rate overshooting conjecture as well as Bjornland's (2009) more recent work solving exchange rate puzzles using the long-run restriction of Blanchard and Quah (1989).

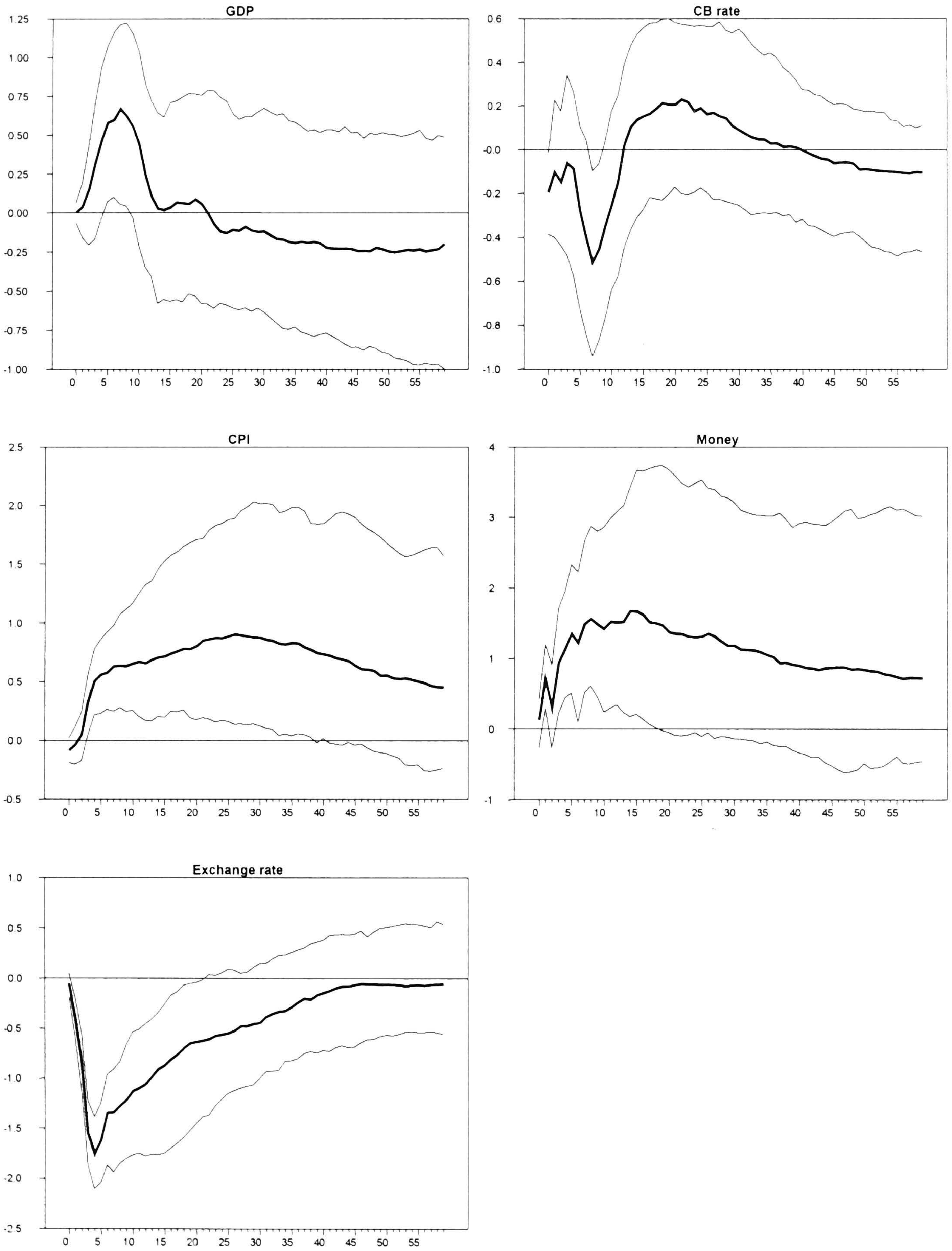
Figure 2.3 The impulse responses to a monetary policy shock



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

Figure 2.4 illustrates the impulse responses to an oil price shock. The price level rises rapidly within two to three months while output tends to increase. Surprisingly, monetary policy seems to react to oil shocks by a reduction in the interest rate. Some studies, such as Bernanke, Gertler and Watson (1997), have found that monetary policy is systematically tightened in response to a rise in the oil price.

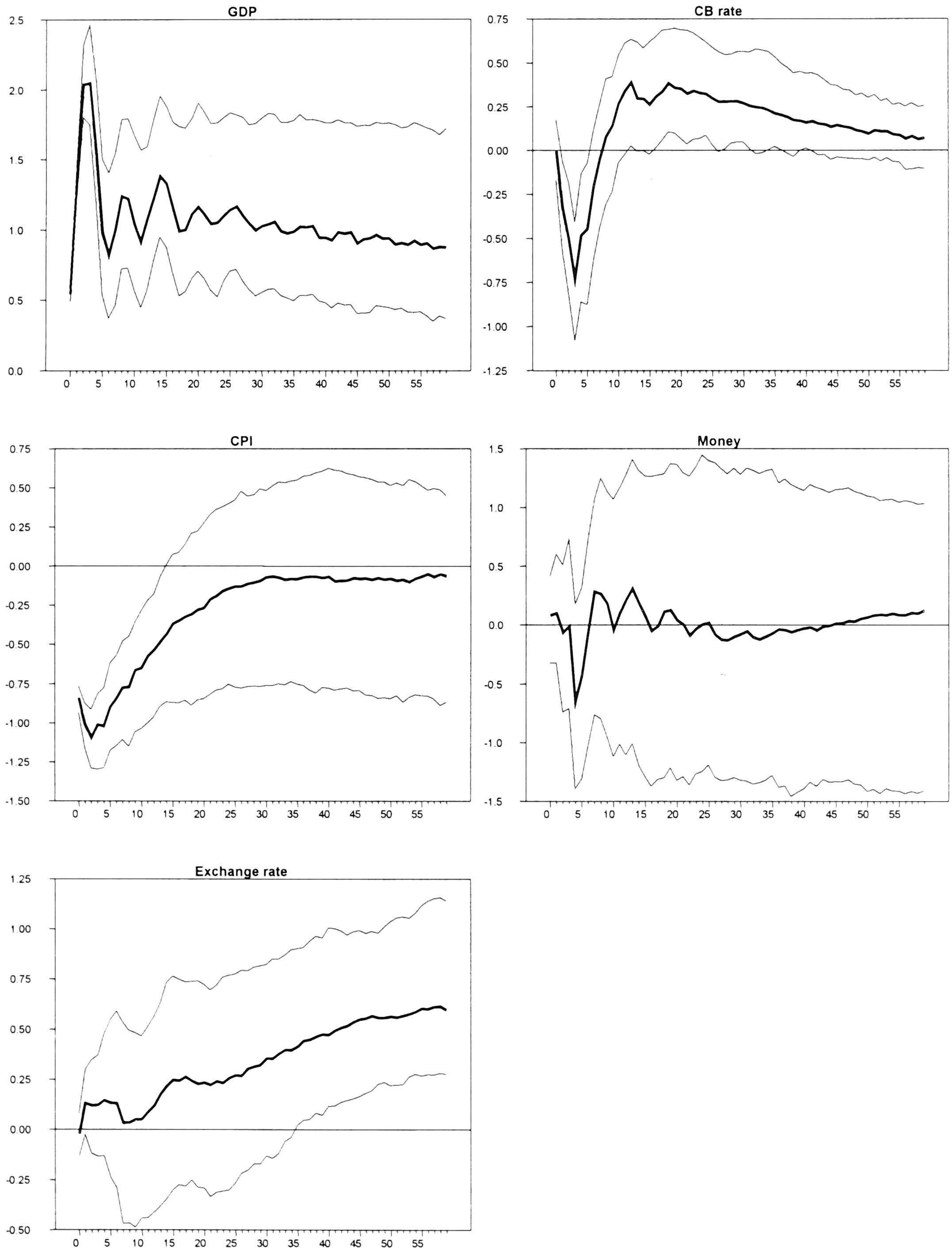
Figure 2.4 The impulse responses to an oil price shock



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

Figure 2.5 provides the impulse responses to an aggregate supply shock. A positive supply shock rapidly raises GDP by about 2 percent and reduces the price level by about 1 percent in two to three months. Its long-run effect on GDP appears to be permanent, stabilizing at around 1 percent. In response to a favourable supply shock, the central bank reacts almost instantly by loosening its monetary policy. The exchange rate responds gradually to depreciate with a peak effect of 0.6 percent.

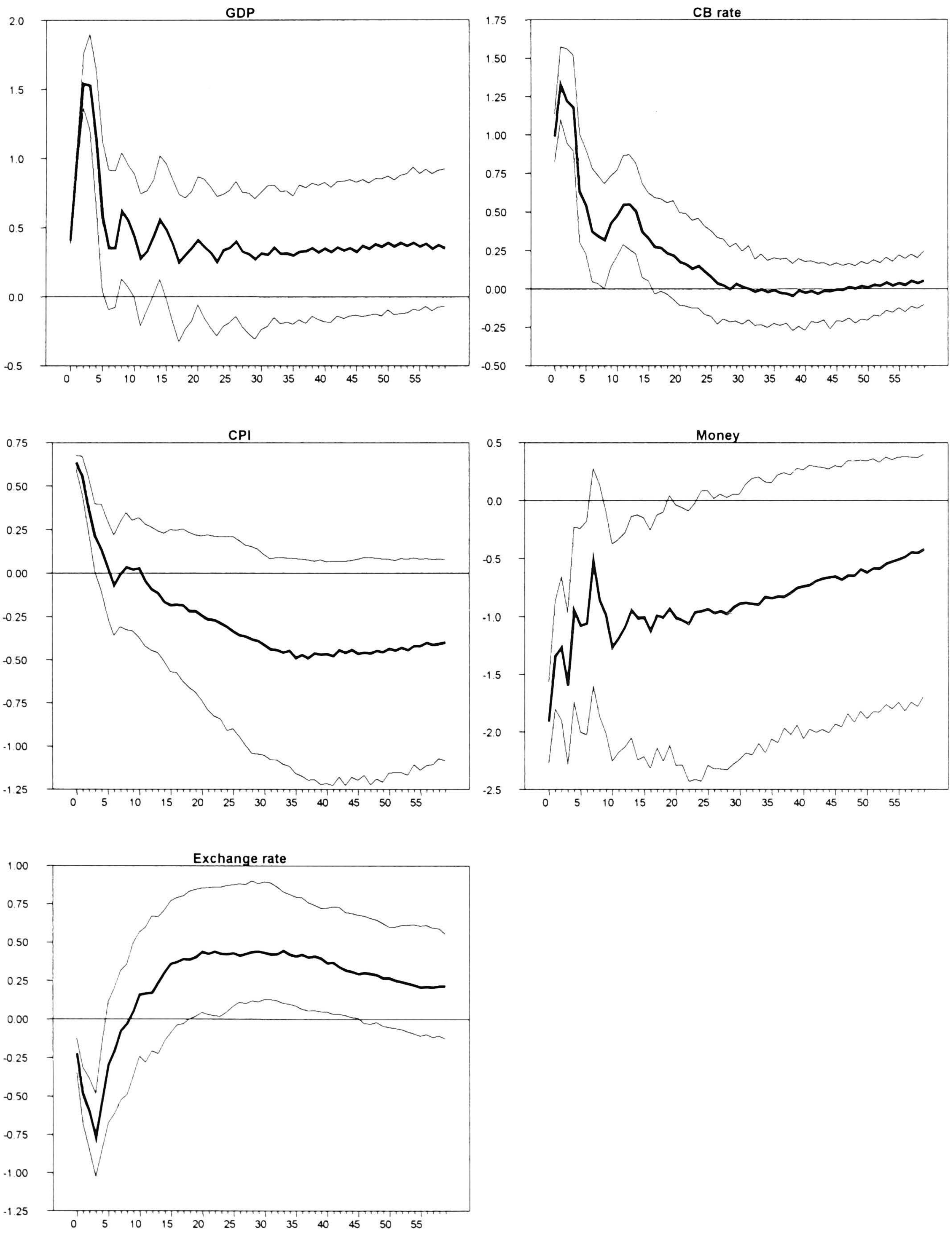
Figure 2.5 The impulse responses to an aggregate supply shock



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

Figure 2.6 shows the impulse response to an aggregate demand shock. A positive demand shock raises both the price level and output which in turn causes the interest rate to increase, money supply to decrease and the exchange rate to appreciate.

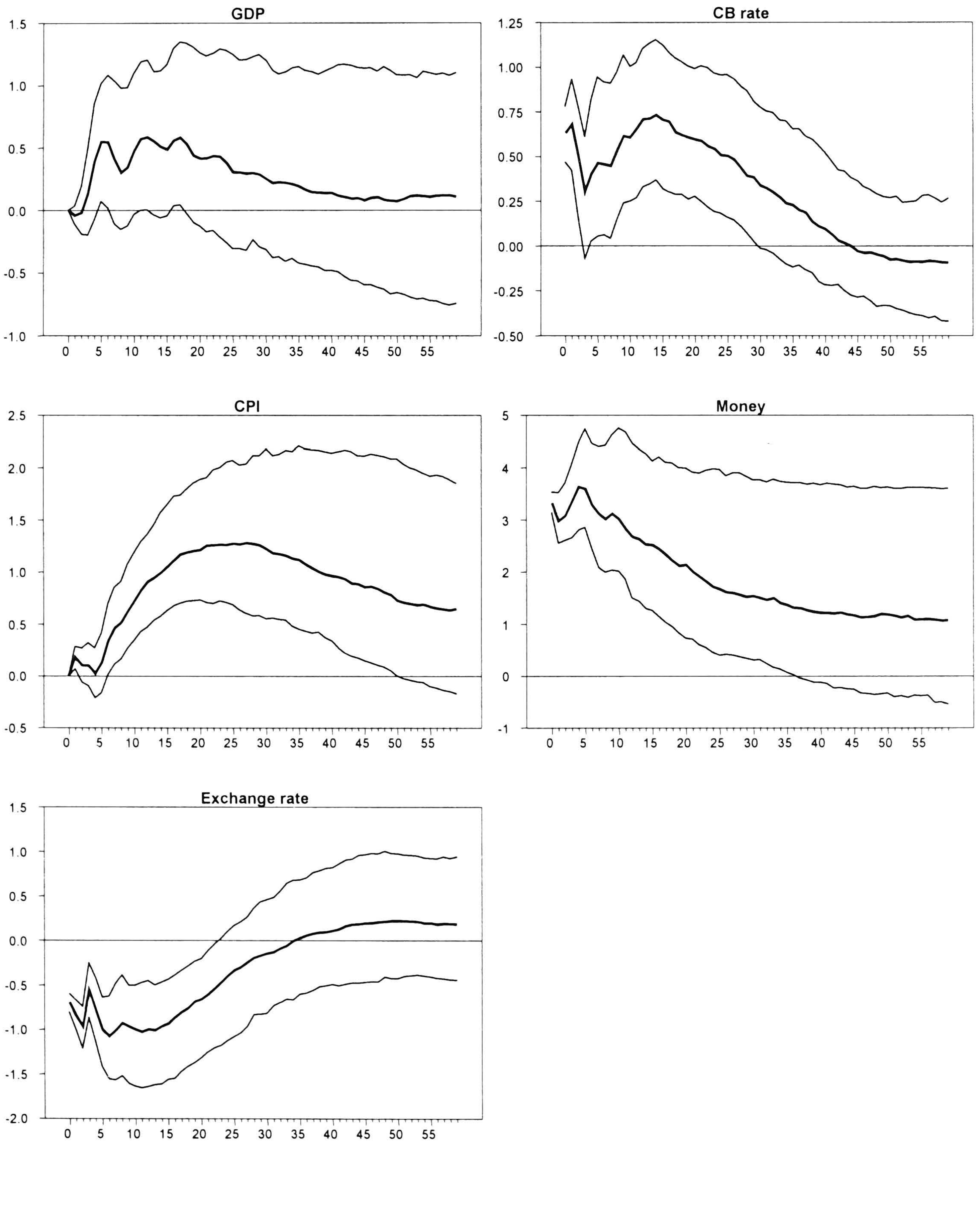
Figure 2.6 The impulse responses to an aggregate demand shock



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

Figure 2.7 demonstrates the impulse response to an LM shock. A positive LM shock raises both the price level and output which in turn causes the interest rate to increase and the exchange rate to appreciate. The effect of LM shocks on the CPI is gradual and persistent with its peak effect occurring after two years which is consistent with the findings of Favara and Giordani (2009).

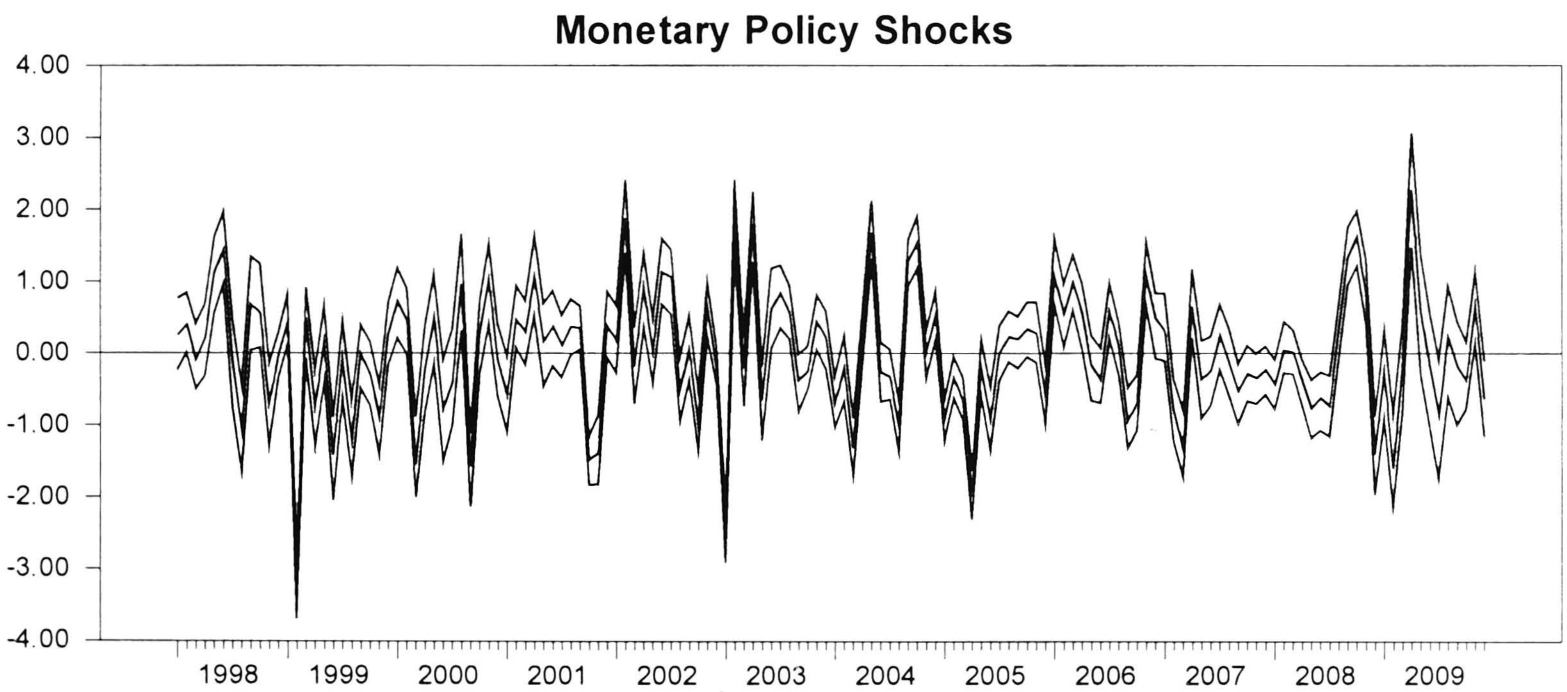
Figure 2.7 The impulse responses to an LM shock



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

It is interesting to know to what extent monetary policy shocks contribute to output and price. The fraction of the variance of the k-step ahead forecast of each variable explained by monetary policy shocks is shown in Figure 2.9. According to median estimates, monetary policy shocks account for around 20 percent of the variation in real GDP, prices, and money at all horizons. For interest rates, 26 percent of the variation explained by the shocks is over the short horizon. The smallest fraction of the variation is explained for the exchange rates. This is surprising, since economists have long suspected that monetary policy shocks are the main source of exchange rate volatility.

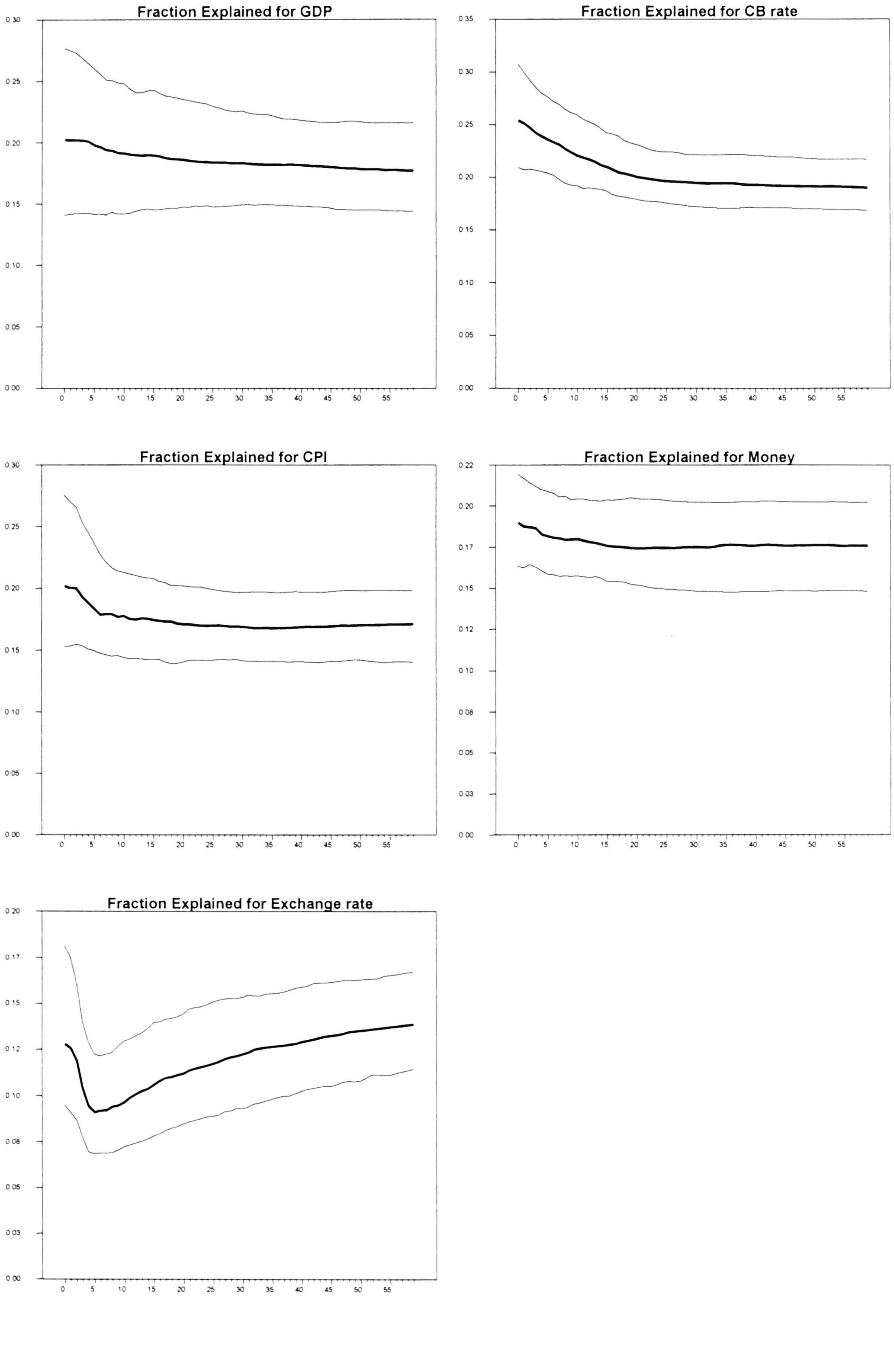
Figure 2.8 The monetary policy shocks identified by sign-restricted SVAR



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of shocks. Thin lines represent 68 error band.

The identified monetary policy shocks are plotted in Figure 2.8. Negative values indicate a loose monetary policy stance while positive values indicate a tight monetary policy stance. For example, if we focus on the last three-year identification results, on average the monetary policy stance was loose except for September to November 2008 and April 2009. For the period covered in the paper, the tightest monetary policy stance was in April 2009. Other identified structural shocks such as aggregate supply and demand shocks, oil price shocks and LM shocks are plotted in Figure 2.17a of the appendix.

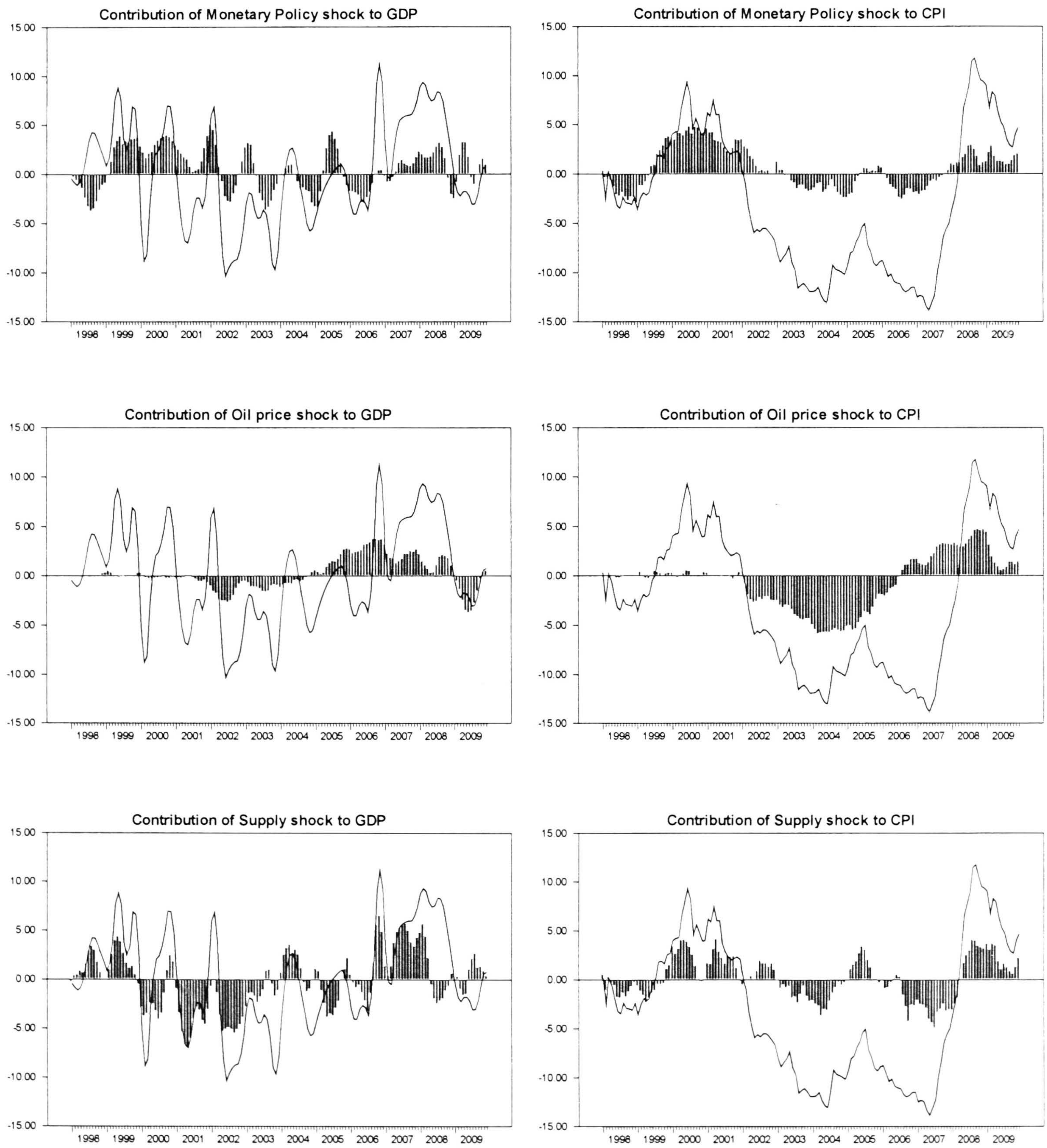
Figure 2.9 Fraction of variance explained by monetary policy shock



Source: Author's calculations
Note: Thin lines represent 68 error band.

The historical decomposition of the real GDP and CPI is displayed in Figures 2.10 and 2.11. These figures emphasize the contribution of each structural shock to a deviation of variables from the baseline at each point in time. The thin line denotes the deviation of the real GDP and price from baseline. The bars (red) denote the component of variables accounted for by each structural shock.

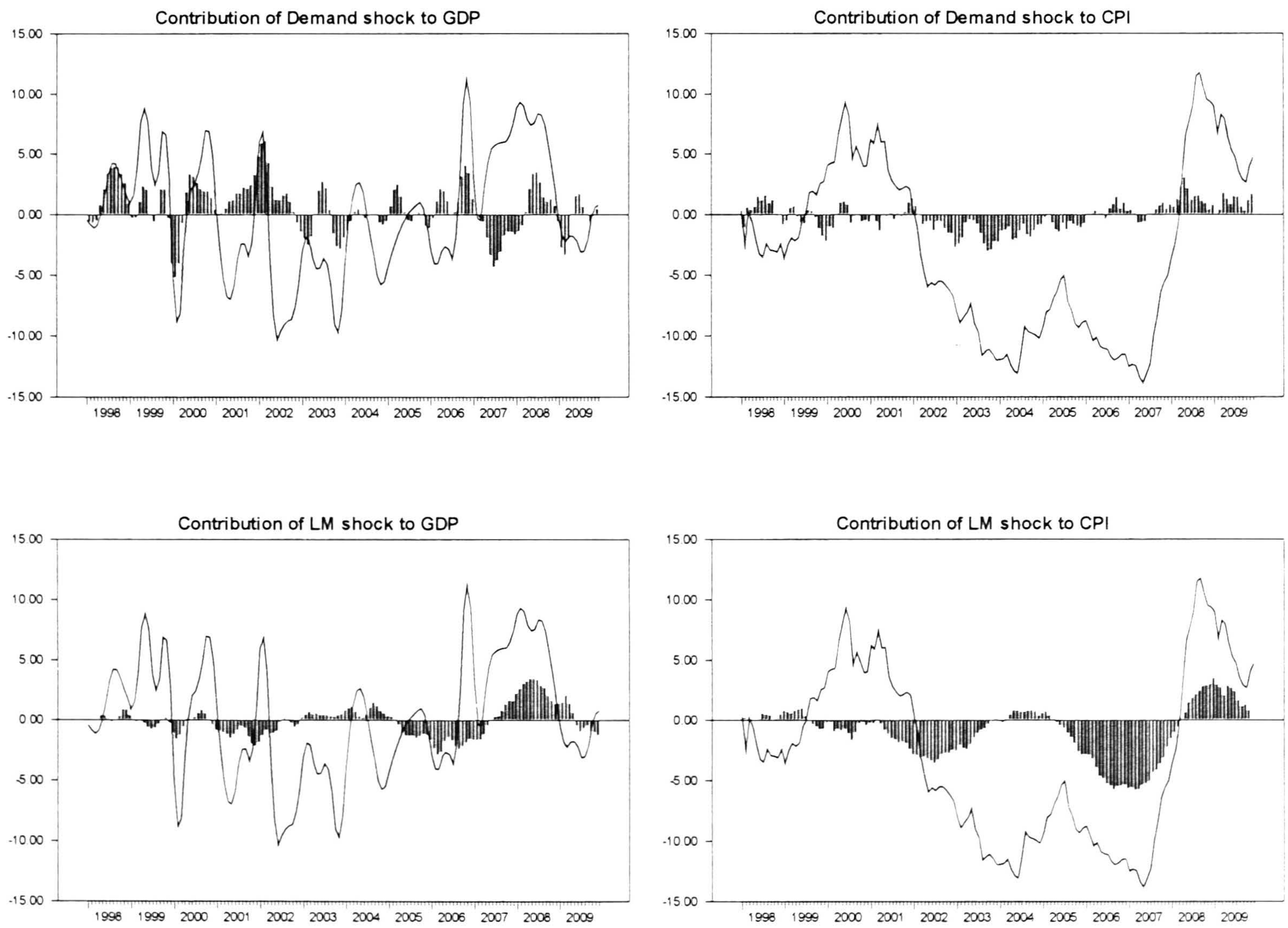
Figure 2.10 Historical decomposition: monetary policy shock, oil shock and supply shock



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables from the baseline.

Besides monetary policy shocks, aggregate supply shocks play an important role in explaining output fluctuation, while oil price shocks and LM shocks play an important role in explaining inflation. For expository convenience, the sample period is divided into four sub-periods for CPI. With regards to the explanation of inflation, monetary policy shocks dominate all other shock for 1998-2001. For 2002- 2005 the decomposition assigns a large role to oil shocks and for 2006-2007, to LM shocks. For 2008-2009 the decomposition suggests that all structural shocks are to be held responsible for increased inflation.

Figure 2.11 Historical decomposition: demand shock and LM shock

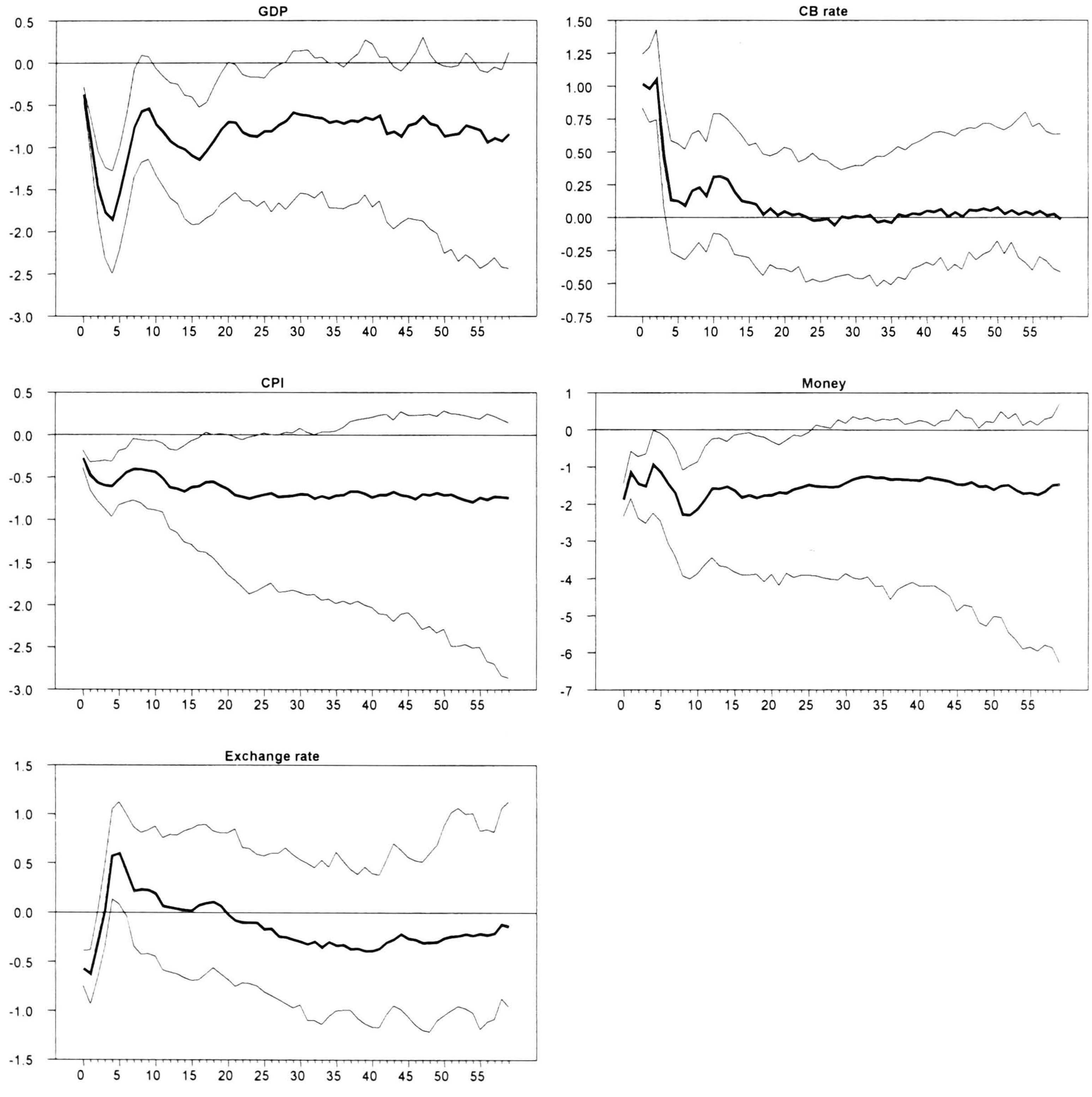


Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables from the baseline.

2.4 Robustness

In this section we carry out an analysis of the robustness of our results. In particular, we checked for robustness to changes in lag length, deterministic terms, and the horizon K for the sign restriction approach. The SVAR model is re-estimated with lag lengths of 3 and 9, with and without constant terms and trend, and with the horizon of K=3, 6, 9. We find that the main qualitative results of section 2.3 are robust to these changes. The SVAR model is also re-estimated with short sample size from 2000:12 to 2009:12 (see Figure 2.12) It seems that no structural break occurred during our sample period.

Figure 2.12 The impulse responses to a monetary policy shock (sample: 2000-2009)

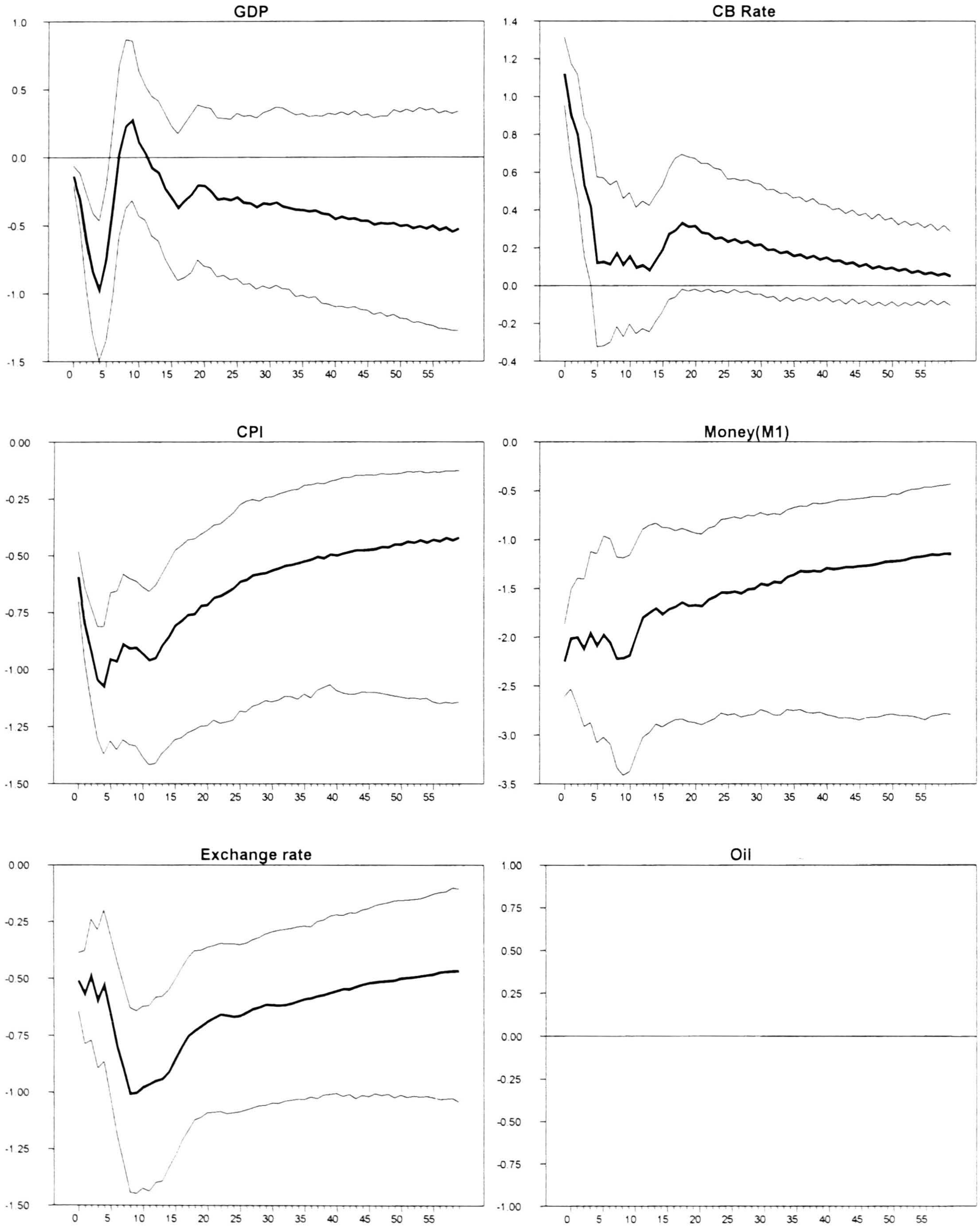


Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

The sign-restricted SVAR models analysed so far have the same set of variables on the right hand sides of each equation. For a more realistic case, we should not allow feedback from foreign exogenous variables such as oil price to domestic variables. In order to impose these restrictions on the VAR model we use a near-VAR model that sets up the equations individually. In particular, the oil price equation is set up having no contemporaneous and lagged domestic variables on the right hand side. However, we cannot use simple Monte Carlo techniques for drawing the posterior to obtain confidence bands (see Doan, 2010). Instead, we need to adopt the Gibbs sampling method.

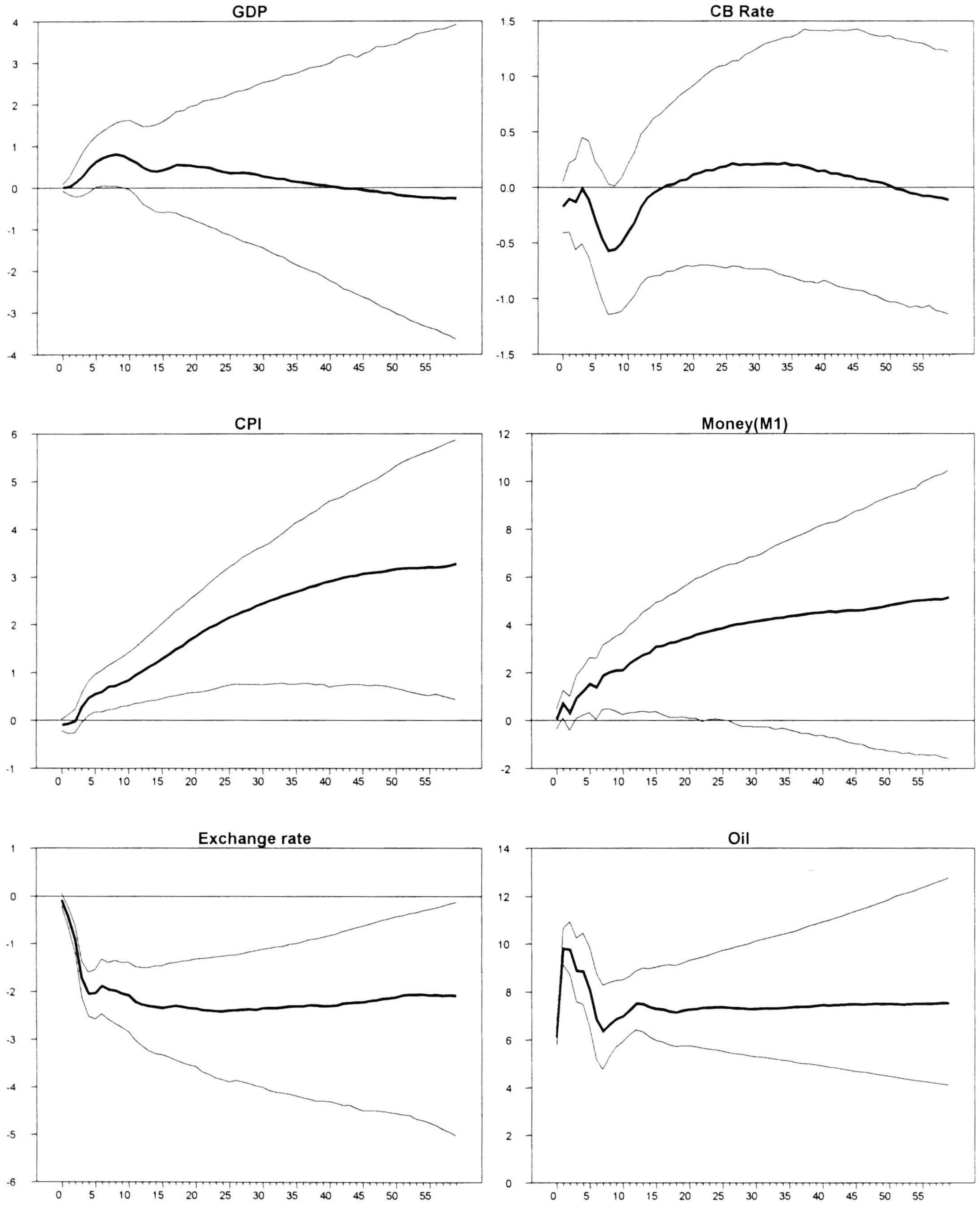
The results from the sign-restricted near-VAR model are shown in Figures 2.13 and 2.14. The impulse responses to monetary policy shocks estimated by the near-VAR model are very similar to those in Figure 2.3. One slight difference is that the exchange rate overshoots its long-run rate with effects delayed three to four months. The impulse responses to oil price shocks are also similar to the corresponding Figure 2.4. The pass-through effect on inflation measured by CPI is more pronounced. The impulse responses to aggregate supply and demand are not shown because they are almost the same as the corresponding figures in the previous section.

Figure 2.13 The impulse responses to a monetary policy shock (near-VAR)



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

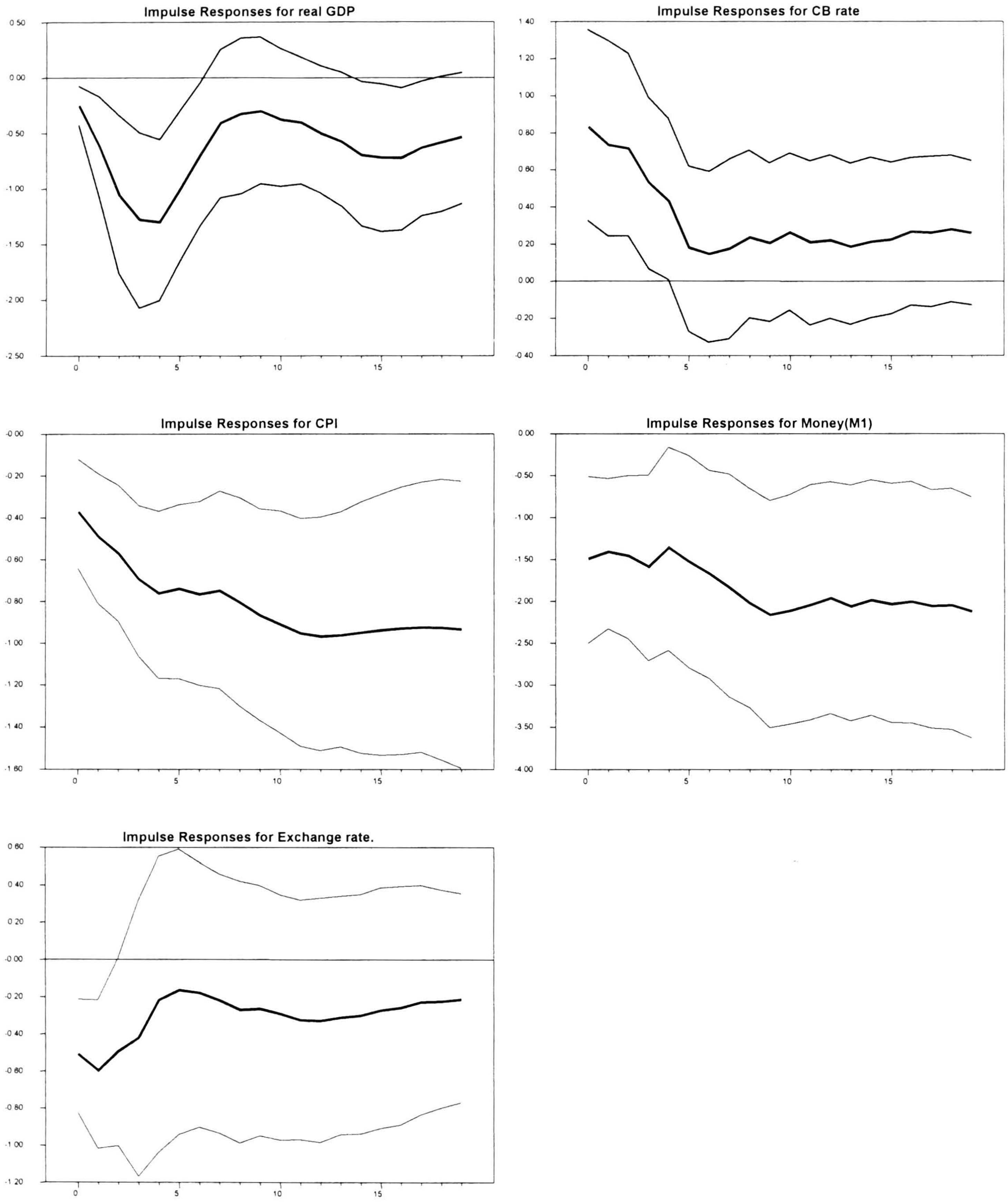
Figure 2.14 The impulse responses to an oil shock (near-VAR)



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

While the penalty function approach used in the previous section does not suffer from the multiple model problems raised by Fry and Pagan (2011), it is somewhat more restrictive than the pure sign restriction approach. Thus, it is interesting to check whether our results are sensitive to the pure sign restriction approach. The impulse responses to monetary policy shocks are robust to the latter approach (Figure 2.15).

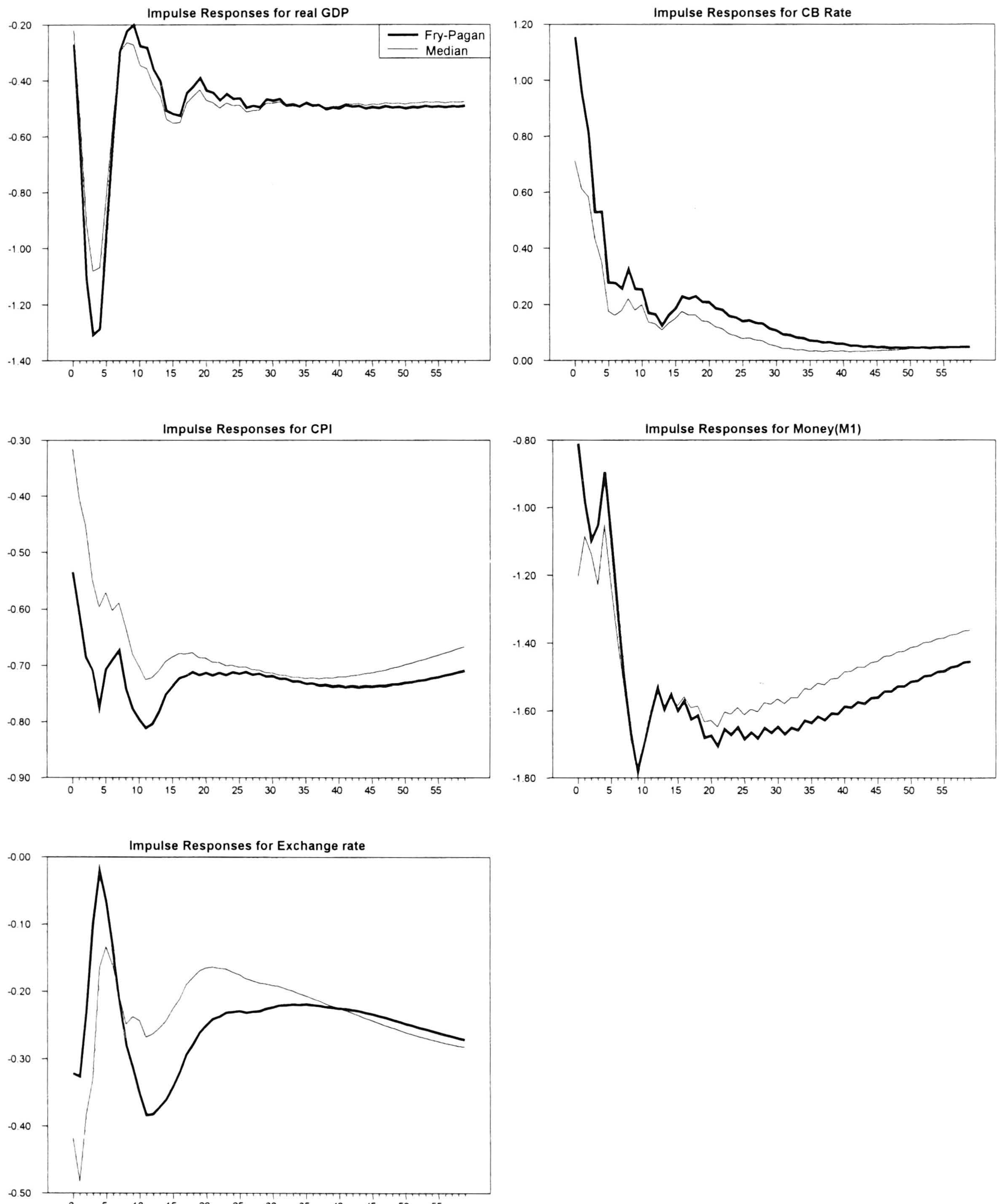
Figure 2.15 The impulse responses to a monetary policy shock (pure sign restriction)



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables. Thin lines represent 68 error band.

Since the pure sign restriction approach is vulnerable to multiple model problems we estimate a median target (MT) and compare it with the median estimates. The MT solution is to choose a single model by minimizing a criterion that provides impulses that are as close to the median as possible. Figure 2.16 shows the median impulses and those estimated from the MT approach. We found that applying the MT method produces little difference except in the magnitudes of the initial impacts of a monetary policy shock.

Figure 2.16 The impulse responses to a monetary policy shock (MT approach)



Source: Author's calculations
Note: Vertical axis scales represent percent deviation of variables.

2.5 Conclusions

This paper attempts to measure the lagged effect of the monetary transmission mechanism on inflation and output in Mongolia using a sign-restricted structural VAR. Compared with the traditional recursive approach, a sign-restricted VAR provides more significant estimation results, consistent with theoretical expectations.

We find the following: first, the lag of the monetary transmission mechanism is about 4 to 12 months for Mongolia. Contractionary monetary policy reduces real GDP by about 1.5 percent after just four months. The price level measured by the CPI reacts relatively slowly with prices dropping by 1.1 percent within one year. Second, monetary policy shocks play a modest role in explaining output and inflation fluctuations. Third, the exchange rate immediately overshoots its long-run equilibrium rate in response to a monetary policy shock, a finding consistent with Dornbusch's (1976) exchange rate overshooting hypothesis. However, according to variance decomposition analysis it is difficult to explain exchange rate volatility by monetary policy shocks. Fourth, the historical decomposition analysis suggests that besides monetary policy shocks, output fluctuations are largely driven by aggregate supply shocks, while inflation is largely driven by oil price and LM shocks.

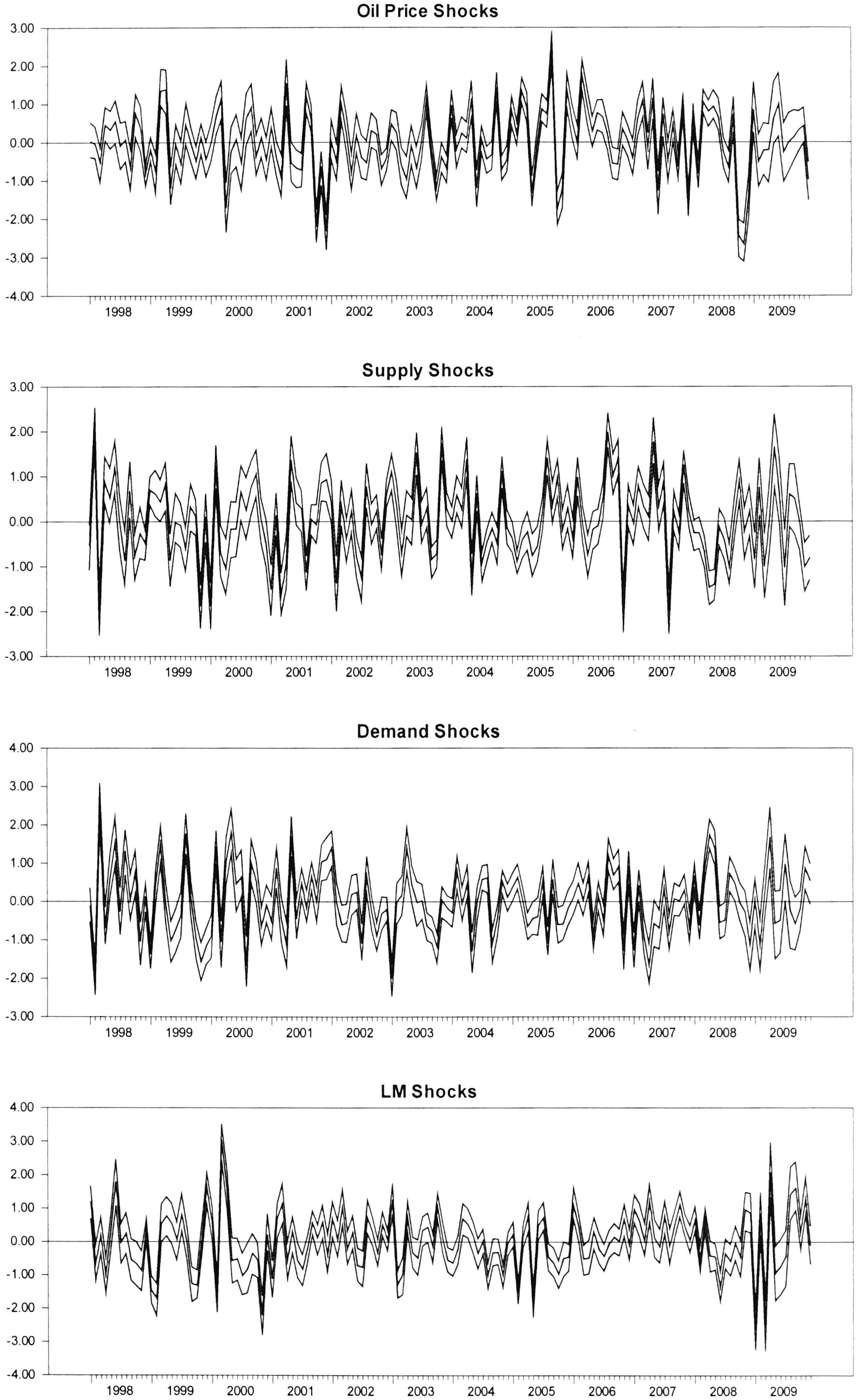
2.6 Appendix

Table 2.2a VAR lag order selection

Lag	LogL	LR	FPE	AIC	SC
1	1373.061	NA	5.73e-16	-18.06839	-17.33933
2	1474.760	186.9074	2.36e-16	-18.95622	-17.49811*
3	1540.043	114.6857	1.60e-16	-19.35193	-17.16477
4	1580.694	68.11796	1.52e-16	-19.41478	-16.49857
5	1639.597	93.92696	1.13e-16	-19.72429	-16.07903
6	1694.040	82.40033	9.05e-17*	-19.97352*	-15.59921
7	1714.666	29.54564	1.16e-16	-19.76576	-14.66240
8	1736.730	29.81539	1.47e-16	-19.57743	-13.74502
9	1789.457	66.97826*	1.25e-16	-19.80348	-13.24201

Note: * indicates lag order selected by the criterion; LR: sequential modified LR test statistic (each test at 5% level); FPE: Final prediction error; AIC: Akaike information criterion; SC: Schwarz information criterion
Source: Author's calculations

Figure 2.17a Oil price, supply, demand and LM shocks identified by sign-restricted SVAR



Chapter 3 INFLATION DYNAMICS IN MONGOLIA

3.1 Introduction

In recent years Mongolia has experienced a rapid growth in money supply while inflation has remained relatively low. From 1999-2004 the Mongolian inflation rate averaged around 6 percent a year, then, following this period, the average rate of inflation almost doubled. There is, however, a debate in the academic literature as to the exact role of money in determining inflation (Woodford, 2008) and other factors such as import prices, especially for petroleum and government spending (e.g. wages, pensions and social transfers) may also have a potential inflationary impact. In Mongolia, the average government wage was below the national average for many years, but since 2006, the trend has reversed. In contrast, exchange rate appreciation due to commodity price hikes could mitigate inflationary pressure. This uncertainty over the causes of inflation makes understanding the factors behind its recent increase in Mongolia an important issue for policy makers.

In order to better understand inflation dynamics and the policy that a central bank implements, this paper develops an empirical model for inflation in Mongolia. In particular, we first estimate long-run markup and money demand relationships using the cointegration procedures, and then construct a single-equation error correction model of inflation with possible nonlinearity.

This paper attempts to address the following questions using both classical and Bayesian approaches. First, what are the main determinants of inflation in the long and short run, and are they stable? Second, does money matter for inflation in the long and short run? Third, what are the driving forces behind the recent continued increase in inflation?

The main findings of the paper are summarized as follows. First, the main determinant of inflation is the markup, capturing impact from unit labor costs, petroleum prices, import prices for Mongolia's trading partners, and the exchange rate. Second, money matters for inflation in both the short and long run: excess narrow-money supply in the money market seems to cause inflation in the long run but adjustment to disequilibria is slow. Third, sustained increases in wages together with petroleum price shocks explain high and volatile inflation in recent years. We also find two inflationary regimes, characterized by the degree of inflation persistence.

The paper consists of six sections. Section 3.2 describes a cointegrated VAR model and section 3.3 explains the data and integration order. Empirical results of the classical approach are presented in section 3.4 and empirical results of the Bayesian approach in section 3.5. Finally, section 3.6 concludes. We model the possible nonlinearity of inflation persistence by the Markov switching model and the results are shown in the appendix.

3.2 Econometric model

Many economic factors can affect inflation. These include: demand side factors that may cause demand pull inflation; monetary factors; and supply side factors that come from the cost push or markup relationships. In order to model the role played by these factors we will apply the method developed by Juselius (1992) and Hendry (2000). In particular, we first estimate long-run markup and money demand relationships using the cointegration procedures and then construct a single-equation error correction model of inflation.

Let x_t be a p-dimensional vector represented by a cointegrated VAR model with r stationary cointegrated vectors

$$\Delta x_t = \Phi D_t + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \alpha\beta' x_{t-1} + \varepsilon_t, \quad t = 1,...,T. \tag{2.1}$$

The residual vector ε_t is assumed to be i.i.d. $N_p(0,\Omega)$, where Ω is positive definite. The model parameters are $\alpha(p \times r)$, $\beta(p \times r)$, $\Gamma_1,...,\Gamma_{k-1}(p \times p)$, and $\Phi(p \times d)$ for some $r \in \{0,1,...,p\}$. Vector $D_t(d \times 1)$ is a constant, trend, seasonal dummies, or other deterministic or exogenous variables. A more compact form of the model can be written as

$$Z_0 = \Phi D + \Gamma Z_2 + \alpha\beta' Z_1 + \varepsilon, \tag{2.2}$$

where $\Gamma = [\Gamma_1 \cdots \Gamma_{k-1}]$, $Z_0 = [\Delta x_1 \cdots \Delta x_T]$ is a $p \times T$ matrix, $Z_1 = [x_{-1} \cdots x_{T-1}]$ is a $p \times T$ matrix, $Z_2 = [Z_{21} \cdots Z_{2T}]$ is a $p(k-1) \times T$ matrix, $Z_{2t} = [\Delta x'_{t-1} \cdots \Delta x'_{t-(k-1)}]'$ is a $p(k-1) \times 1$ vector, $D = [D_1 \cdots D_T]$ is a $d \times T$ matrix and $\varepsilon = [\varepsilon_1 \cdots \varepsilon_T]$ is a $p \times T$ matrix. We use the notation $\mathfrak{D} = \{Z_0, Z_1, Z_2, D\}$ for all available data.

The approach by Johansen (1988) estimates the system (2.1) maximum likelihood by imposing the restriction $\Pi = \alpha\beta'$ for a given value of *r*. It can be shown that the maximum likelihood estimate for β equals the matrix containing *r* eigenvectors corresponding to the *r* largest eigenvalues. We can use the eigenvalues, say $\lambda_1 > \lambda_2 > \ldots > \lambda_p$, to test the hypothesis of rank of matrix Π. This is the so-called trace test:

$$\lambda_{trace}(r_0) = -T \sum_{i=r_0+1}^{p} \log(1-\lambda_i). \tag{2.3}$$

It checks whether the smallest $p - r_0$ eigenvalues are significantly different from zero.

3.3 Data and Integration

Money demand and markup equations were estimated using quarterly data from 1997:4 to 2009:4. We estimated the money demand function using four variables: output, consumer price index, narrow money (M1) and the banks' time deposit interest rate. The consumer price index and real GDP (in constant price of 2005) are taken from the bulletin of the National Statistical Office of Mongolia. Data for narrow money and banks' deposit rate are from the bulletin of the Bank of Mongolia. We also estimated the price equation using four variables: consumer price index, unit labor cost, petroleum price index in domestic currency and import price. Unit labor cost is calculated using average monthly wage, employment and output data. For import prices, we used the consumer price index of China as proxy variables because China is the main trading partner of Mongolia. Average monthly wage, employment and petroleum price data were taken from the bulletin and yearbook of the National Statistical Office of Mongolia. The consumer price index of China is taken from the OECD statistics website. All variables except the interest rate are in logarithms and are seasonally adjusted using the Census X12 approach. The augmented Dickey-Fuller (ADF) test is used to determine the orders of integration for variables (see Table 3.14a of the appendix). All variables appear to be integrated of order one.

3.4 Estimation results of the Classical approach

We model inflation dynamics in Mongolia combining two main theories that operate through markup and money demand. From a general-to-specific point of view, estimating a vector autoregressive model that includes all variables, related markup and money demand would be more appropriate. However, because of the small sample size, we did not employ such a strategy. Instead, we first estimated long-run relationships from two separate VARs and then modelled short-run inflation dynamics with error correction terms that incorporate feedback from both relationships.

A. Long-run relationships

Money demand

We estimated long-run money demand using a standard specification of monetary theory:

$$\frac{M^d}{P} = f(Y, R) \tag{4.1}$$

where M^d is the demand for monetary aggregate, P is the consumer price index, Y is real level of economic output, and R is the rate of return on alternative assets. In log-linear form it can be written as follows[3]:

$$m1 - p = \beta_0 + \beta_1 y + \beta_2 i \tag{4.2}$$

where $m1$ is narrow money, p consumer price index, y real GDP and i is interest rate of bank deposit in domestic currency.

The lag length of the VAR is chosen according to Schwarz and Hannan-Quinn information criteria, which select one lag (see Table 3.15a in the appendix). In order to check the assumptions underlying the model, several misspecification tests were conducted (see Table 3.16a in the appendix). A LM test indicated that the null hypothesis cannot be rejected against the alternative hypothesis of first and fourth order autocorrelation. According to a Doornik-Hansen normality test, the null hypothesis that residuals have skewness of zero and kurtosis of 3 is not rejected, while a multivariate ARCH test is rejected. Only residuals from the deposit rate equation show a significant ARCH effect. However, Rahbek, Hansen and Dennis (2002) have proven that the Johansen test is robust to moderate residual with ARCH, therefore one lag may be sufficient for modelling the dynamics of the variables.

Table 3.1 shows the results of Johansen's multivariate cointegration test on the real narrow money M1. Following Bruggeman, Donati and Warne (2003), a Bartlett

[3] We estimated other opportunity cost variables such as exchange rate and inflation rate for money demand, but these variables were insignificant. We also could not find meaningful and significant money demand equation for broad money (M2).

correction for the trace test (Johansen, 2002) and bootstrap p-values are used. The number of bootstrap replications is 1000. From Table 3.1 we can infer that there is only one cointegration vector of money demand. All coefficients have their expected signs and magnitudes. According to a weak exogeneity test, the null hypothesis of interest rate is rejected. Moreover, the hypothesis of long-run unit income homogeneity is not rejected.

Table 3.1 Cointegration analysis: real money demand

Johansen test			
Eigenvalue	0.5848	0.2009	0.0007
Null hypothesis, H_0	$r = 0$	$r \le 1$	$r \le 2$
LR trace	48.5738	9.8971	0.0316
Asymptotic p-value	0.0001	0.2887	0.8590
Bootstrap p-value	0.0020	0.3423	0.8799
LR trace (Bartlett corrected)	43.0442	8.6226	0.0244
Asymptotic p-value	0.0009	0.4014	0.8758
Bootstrap p-value	0.0010	0.3654	0.8869
Standardized cointegration vector and adjustment coefficients			
Variables	*m*1-*p*	*y*	*i*
Cointegrating vector , β'	1.0000	-1.0046	6.5287
Adjustment coefficients, α	-0.0438	0.0312	-0.0635
Weak exogeneity test			
χ^2	1.0843	0.8648	28.3824
asym. p-value	0.2977	0.3524	0.0000
Hypothesis	χ^2	*deg.free*	*p-value*
$H_0^1: \beta'=(1\ -1\ *)$	0.0004	1	0.9842
$H_0^2: \beta'=(1\ -1\ *)$, $\alpha_{m1}=0$, $\alpha_{gdp}=0$	1.7000	3	0.6369
Restricted cointegrating vector (H_0^2)	*m*1-*p*	*y*	*i*
Cointegrating vector, β'	1.0000	-1.0000	7.0987
standard error	-	-	0.6294
α-coefficient	0.0000	0.0000	-0.0565
standard error	-	-	0.0075

Source: Author's calculation

Next, we will consider the stability of parameters in the VAR model using the three constancy tests. First, the most important is the non-zero eigenvalue fluctuation test studied by Hansen and Johansen (1999). Second, the constancy of β is checked using Nyblom's (1989) test. Finally, we examined the constancy of the Φ, and α parameters using Ploberger, Kramer, and Kontrus (1989).

The results of the constancy analysis are shown in Table 3.2. As can be seen, the eigenvalue, λ_1 appears to be constant over the period 2003:Q2 – 2009:Q4 regardless of whether the Φ parameter is updated or not. The null hypothesis for the constancy of β is not rejected according to the supremum and mean test of Nyblom. Further, the Ploberger-Kramer-Kontrus fluctuation test of the constancy of parameters Φ, α is not rejected for each equation. Based on the results of the constancy test, we conclude that the long-run vector of money demand, β is constant over the sample period.

Table 3.2 Stability of the money demand

1. Hansen-Johansen fluctuation test of the constancy of non-zero eigenvalues			
Updating of $\hat{\Phi}^{(t)}$	*test value*	*asym. p-val*	*boot. p-val*
$\sup_{t\in\mathbb{T}} \tau_{t\mid T}(\lambda_1)$	0.1221	1.0000	0.8468
Conditonal on $\hat{\Phi}^{(T)}$			
$\sup_{t\in\mathbb{T}} \tau_{t\mid T}(\lambda_1)$	0.1254	1.0000	0.7628
2. Nyblom test for the constancy of cointegrating vector β			
Updating of $\hat{\Phi}^{(t)}$	*test value*	*asym. p-val*	*boot. p-val*
$\sup_{t\in\mathbb{T}} Q_T^t(S)$	0.6581	0.7810	0.5816
$mean_{t\in\mathbb{T}} Q_T^t(S)$	0.2867	0.4132	0.3684
Conditonal on $\hat{\Phi}^{(T)}$			
$\sup_{t\in\mathbb{T}} Q_T^t(S)$	0.6502	0.7880	0.4014
$mean_{t\in\mathbb{T}} Q_T^t(S)$	0.2549	0.4730	0.2943
3. Ploberger-Kramer-Kontrus fluctuation test of the constancy of parameters Φ, α			
equation	*S(2)*	*asym. p-val*	*boot. p-val*
*m*1-*p*	0.5753	0.9891	0.6436
y	0.6069	0.9791	0.5916
i	0.6445	0.9606	0.5586

Note: The experiment period is given by $\mathbb{T}$ ={2003:Q2, ... , 2009:Q4}. The model is unrestricted with 1 cointegration relationship
Source: Author's calculations

Markup

In the long run, the domestic price is a markup over total unit cost. Following the paper of De Brouwer and Ericsson (1998) we can write the determinants of the long-run price as follows:

$$P = \mu \cdot \left(ULC^{\gamma}\right)\left(IP^{\delta}\right)\left(PET^{\kappa}\right), \tag{4.3}$$

where P is the consumer price index[4], $\mu - 1$ is the retail markup over costs, ULC is an index of the nominal cost of labour per unit of output, IP is an index of import price in domestic currency, and PET an index of petroleum prices in domestic currency. The elasticities of the CPI with respect to ULC, IP, and PET are γ, δ, and κ respectively.

In practice (4.3) is expressed in the log-linear equation:

$$p = \ln\left(\mu\right) + \gamma \cdot ulc + \delta \cdot ip + \kappa \cdot pet, \tag{4.4}$$

where the logarithms of variables are denoted by lowercase letters. It is typically assumed that linear homogeneity holds for the price markup equation, which means $\gamma + \delta + \kappa = 1$. Under this hypothesis, (4.4) can be rewritten as

$\ln(\mu) = -\left[\gamma(ulc - p) + \delta(ip - p) + \kappa(pet - p)\right]$,where term *(ulc-p)* is real marginal cost. According to Gali and Gertler (1999), Gali, Gertler and Lopez-Salido (2001) and Sbordone (2002), this term is an important element of the New Keynesian Phillips curve instead of the output gap. Furthermore, the purchasing power parity is embedded as the term *(ip-p)* in the markup equation, and the last term is real petrol price in energy markets.

According to the Schwarz and Hannan-Quinn information criteria, the lag length of one quarter may be enough to describe the data generation process (see Table 3.15a).

---

[4] It includes petroleum products. The direct contribution of rise in petroleum prices to inflation is negligible since petroleum has a 1.6 % weight in the CPI.

Several misspecification tests are reported in Table 3.16a in the appendix. The LM tests indicate that the null hypothesis cannot be rejected against the alternative hypothesis of first and fourth order autocorrelation. According to the Doornik-Hansen normality test, the null hypothesis that residuals have skewness of zero and kurtosis of 3 is not rejected. Multivariate ARCH is not rejected for the null hypothesis of no ARCH effect. Thus, one lag seems to be enough for our price model.

Table 3.3 shows the results of the Johansen multivariate cointegration test for the price index, unit labor cost, import price index and petroleum price. Also, following Bruggeman et al. (2003), the Bartlett correction for the trace test (Johansen, 2002) and bootstrap p-values are used. From Table 3.3 we can infer that there is only one cointegration vector of price markup equation, (4.4). All coefficients have their expected signs. The estimated cointegrating vector implies that in the long run, the price level is 33 percent dependent on unit labour cost, 44 percent on import price and 22 percent on petroleum price. The estimated coefficient of the petroleum price implicitly calculates the impact of the exchange rate because the petroleum price is in domestic currency. The sum of the coefficients is approximately one and the hypothesis of long-run unit homogeneity is not rejected. In other words, a 1 percent increase in the price of each input leads to a 1 percent increase in consumer prices. According to a weak exogeneity test, the null hypothesis is rejected for the price and import price.

Table 3.3 Cointegration analysis: Markup

Johansen test				
Eigenvalue	0.5725	0.2208	0.2032	0.0051
Null hypothesis	r = 0	$r \leq 1$	$r \leq 2$	$r \leq 3$
LR trace	62.5870	22.6428	10.9159	0.2401
Asymptotic p-value	0.0012	0.2641	0.2167	0.6241
Bootstrap p-value	0.0060	0.3804	0.3323	0.7217
LR trace (Bartlett corrected)	54.3702	19.1947	7.9564	0.1539
Asymptotic p-value	0.0110	0.4790	0.4701	0.6948
Bootstrap p-value	0.0030	0.3804	0.3894	0.7217
Standardized cointegration vector				
Variables	*p*	*ulc*	*ip*	*pet*
Cointegrating vector , β'	1.0000	-0.3322	-0.4364	-0.2145
Weak exogeneity test				
χ^2 (1)	18.2013	0.6398	7.1183	0.2242
asym. p-value	0.0000	0.4238	0.0076	0.6358
Hypothesis	χ^2	*deg.free*	*p-value*	*boot. p-val*
$H_0 : \gamma + \delta + \kappa = 1$	0.0046	1	0.9465	0.9530
Restricted cointegrating vector (H_0)	*p*	*ulc*	*ip*	*pet*
Cointegrating vector, β'	1	-0.3297	-0.4560	-0.2143
standard error	-	0.0349	0.0128	0.0294

Source: Author's calculations

The results of the constancy analysis are reported in Table 3.4. As can be seen from the table, the eigenvalue, $\lambda 1$, appears to be constant over the period 2002:Q4 – 2009:Q4 regardless of whether the Φ parameter is updated or not. The null hypothesis of the constancy of β is not rejected according to the supremum and mean test of Nyblom. Also, the Ploberger-Kramer-Kontrus fluctuation test of the constancy of parameters Φ, α is not rejected except for the consumer price and import price equation. However, the recursive graph of the test suggests that the parameters of these equations are reasonably constant (see Figure 3.9a in the appendix). In brief, based on the constancy test, we conclude that the long-run vector of price, β is constant over the sample period.

Table 3.4 Stability of the markup equation

1. Hansen-Johansen fluctuation test of the constancy of non-zero eigenvalues			
Updating of $\hat{\Phi}^{(t)}$	*test value*	*asym. p-val*	*boot. p-val*
$\sup_{t\in\mathbb{T}} \tau_{t\mid T}(\lambda_1)$	0.4199	0.9946	0.3834
Conditonal on $\hat{\Phi}^{(T)}$			
$\sup_{t\in\mathbb{T}} \tau_{t\mid T}(\lambda_1)$	0.4610	0.9838	0.1792
2. Nyblom test for the constancy of cointegrating vector β			
Updating of $\hat{\Phi}^{(t)}$	*test value*	*asym. p-val*	*boot. p-val*
$\sup_{t\in\mathbb{T}} Q_T^t(S)$	0.7366	0.8737	0.6907
$mean_{t\in\mathbb{T}} Q_T^t(S)$	0.1883	0.8289	0.8298
Conditonal on $\hat{\Phi}^{(T)}$			
$\sup_{t\in\mathbb{T}} Q_T^t(S)$	0.5111	0.9771	0.8018
$mean_{t\in\mathbb{T}} Q_T^t(S)$	0.1570	0.8948	0.8288
3. Ploberger-Kramer-Kontrus fluctuation test of the constancy of parameters Φ, α			
equation	*S(2)*	*asym. p-val*	*boot. p-val*
p	1.5980	0.0240	0.0360
ulc	1.1881	0.2235	0.1872
ip	2.0375	0.0010	0.0130
pet	0.9404	0.5637	0.4354

Note: The experiment period is given by $\mathbb{T}$ ={2002:Q4, ... , 2009:Q4}. The model is unrestricted with 1 cointegration relationship
Source: Author's calculations

B. Short-run inflation dynamics

According to Hendry (1995), single equation modelling for inflation is risky if the cointegrating vector is estimated jointly with inflation equation dynamics because weak exogeneity tests for the import prices are rejected. However, there are several solutions to this problem. The first approach is to model the CPI and import price as a subsystem, conditional on other variables. The second is to adopt an error correction term from the system and then to estimate a single equation error correction model. This paper uses the second approach. We start with the estimation of the following error correction single equation:

$$\Delta p_t = \varphi_0 + \sum_{i=1}^{3} \varphi_{1i} \Delta p_{t-i} + \sum_{i=0}^{3} \varphi_{2i} \Delta m_{t-i} + \sum_{i=0}^{1} \varphi_{3i} \Delta ulc_{t-i} + \sum_{i=0}^{1} \varphi_{4i} \Delta ip_{t-i}$$
$$+ \sum_{i=0}^{1} \varphi_{5i} \Delta pet_{t-i} + \theta_1 ect_{\inf,t-1} + \theta_2 ect_{mon,t-2} \qquad (4.5)$$

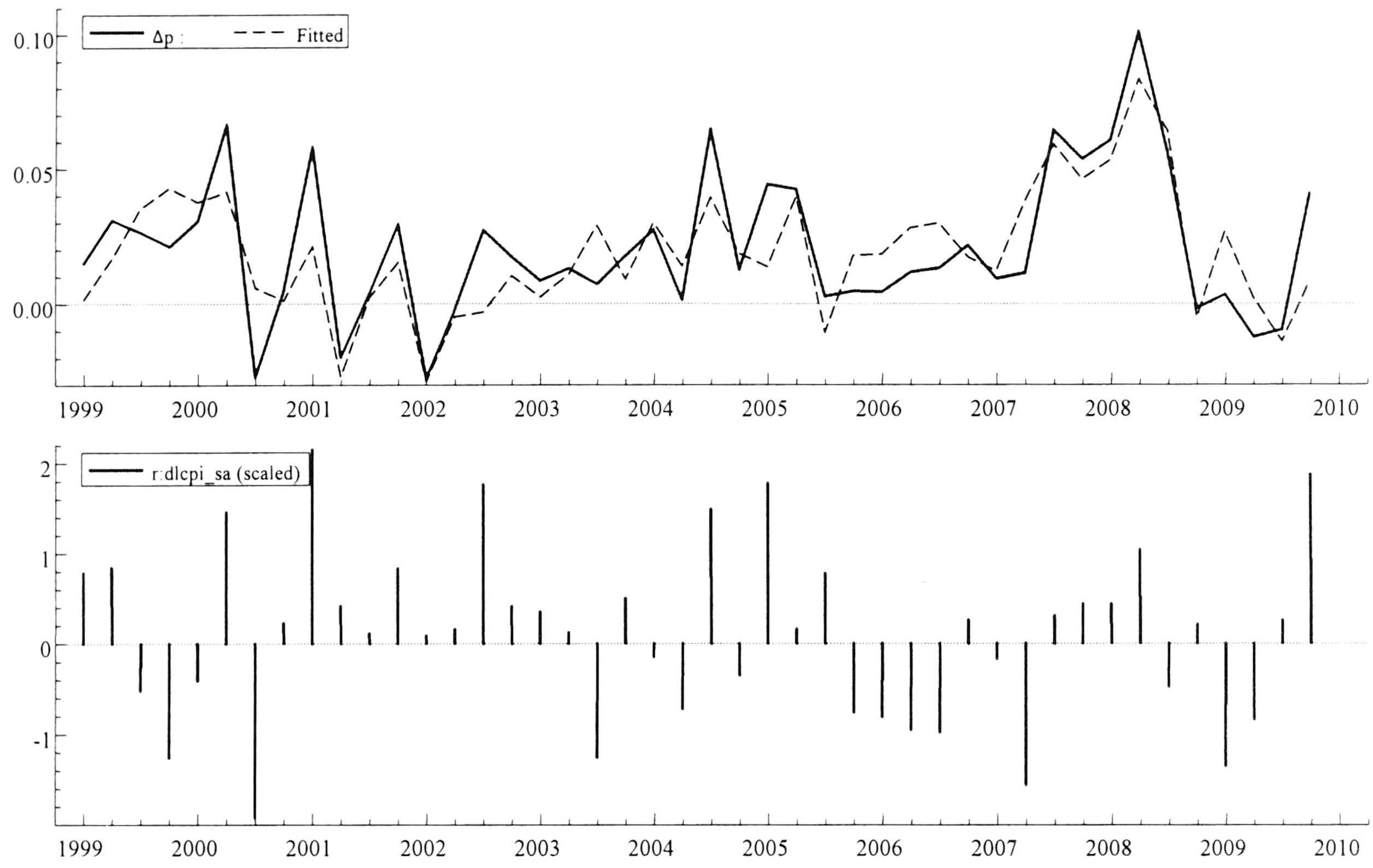
where $ect_{\inf,t-1}$ and $ect_{mon,t-2}$[5] are the error correction terms for the markup and excess money, respectively. Equation (4.5) has both short and long-run elements. The long run elements of the model are characterized by the two error correction terms which show the amount of disequilibrium transmitted in each period into the inflation. The short run elements of the model are accounted for by the inclusion of variables in first differences.

Table 3.5 shows the regression results with a selected specific parsimonious model. Figure 3.1 shows actual, fitted values, and residuals of the parsimonious model. The estimated model fits the data reasonably well and indicates no large residuals.

General and specific models do not indicate any problem from misspecification tests. The instability test is a joint parameter-constancy and variance-change test from Hansen (1992) and is insignificant. Figure 3.2 plots the recursive coefficient estimates, the one-step residuals with ± 2 standard errors and the one-step, break-point and forecast Chow test. Most recursively estimated parameters are relatively stable, with downward drift in the coefficient of inflation inertia. Further, all constancy tests are insignificant. For the specific model, all variables have the expected sign and are highly significant. The constant was insignificant, so there is no evidence of autonomous inflation. The error correction model suggests that inflation does not significantly respond to excess supply of money (M1), which is consistent with the literature. For instance, Durevall and Ndung'u (1999), Hendry (2000) and Kuijs (2002) did not find evidence that the excess money causes inflation.

[5] Lag of 1 for error correction term for excess money was insignificant and the expected sign was wrong.

Figure 3.1 Inflation model fit and residuals



Source: Author's calculations

Table 3.5 General to specific model

1. General model

Variables	Coefficient	Standard error	t-value	t-prob
constant	0.0196	0.0100	1.96	0.06
Δp_{t-1}	-0.1251	0.1323	-0.95	0.35
Δp_{t-2}	-0.1511	0.1200	-1.26	0.22
Δp_{t-3}	0.2632	0.1281	2.05	0.05
Δulc_{t}	0.0048	0.0436	0.11	0.91
Δulc_{t-1}	-0.0172	0.0533	-0.32	0.75
Δip_{t}	-0.4665	0.5083	-0.92	0.37
Δip_{t-1}	1.1417	0.5366	2.13	0.04
Δpet_{t}	0.1057	0.0375	2.82	0.01
Δpet_{t-1}	0.0192	0.0384	0.50	0.62
$\Delta m1_{t}$	-0.0561	0.0485	-1.16	0.26
$\Delta m1_{t-1}$	-0.0004	0.0460	-0.01	0.99
$\Delta m1_{t-2}$	0.0485	0.0654	0.74	0.47
$\Delta m1_{t-3}$	-0.0401	0.0637	-0.63	0.53
$ECT_{inf,t-1}$	-0.2997	0.1348	-2.22	0.03
$ECT_{mon,t-2}$	0.0353	0.0299	1.18	0.25

Misspecification test		*value*	*p-value*
AR 1-4 test	F(4, 24)	0.7539	0.5653
ARCH 1-4 test	F(4, 20)	0.5862	0.6763
Normality test	$\chi^2(2)$	0.6074	0.7381
Hetero test	χ^2 (30)	27.9860	0.5712
RESET test	F(1, 5)	0.7329	0.3995

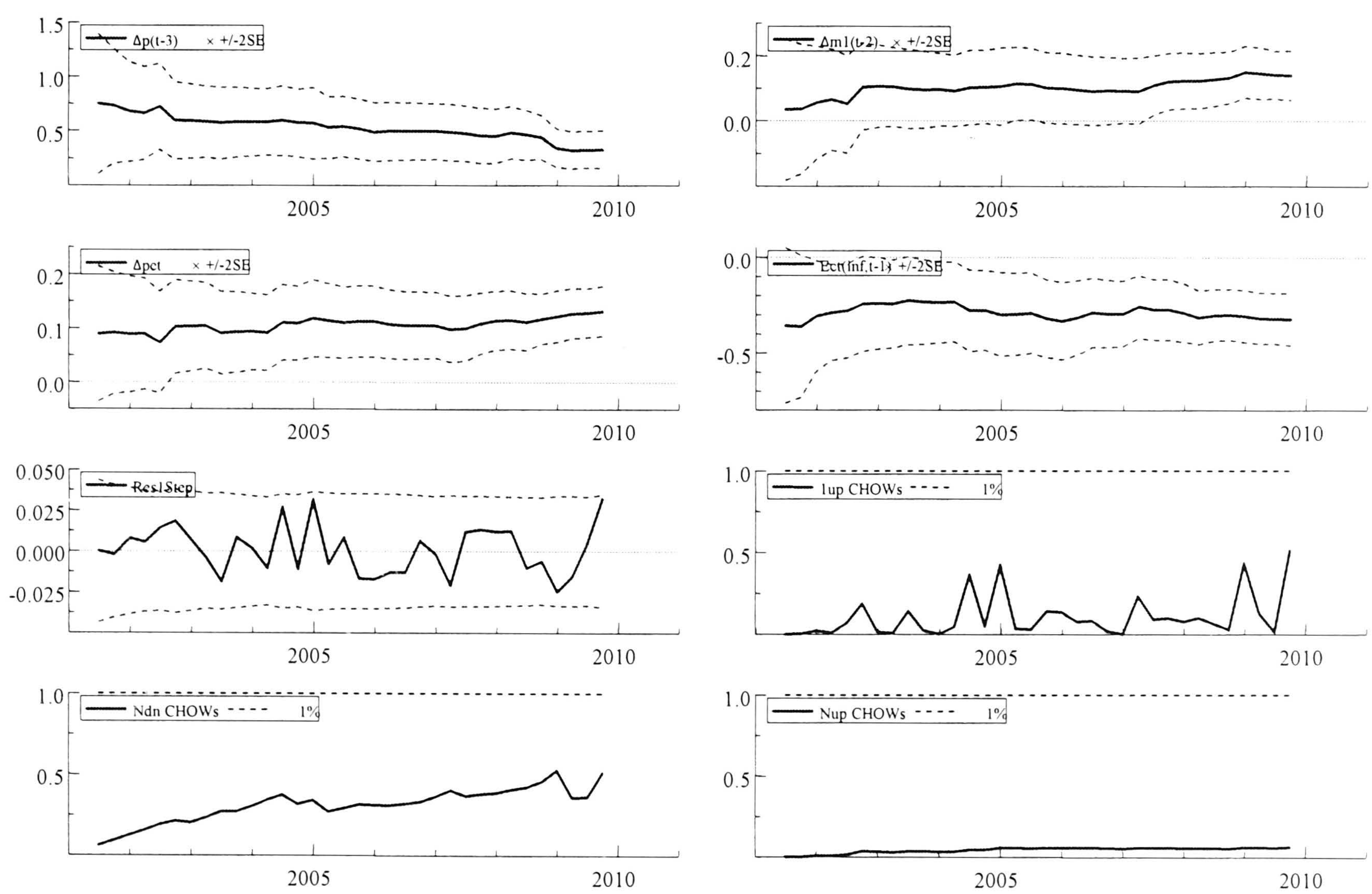
2. Specific model

Variables	Coefficient	Standard error	t-value	t-prob
Δp_{t-3}	0.3275	0.0854	3.83	0.0004
$\Delta m1_{t-2}$	0.1398	0.0380	3.68	0.0007
Δpet_t	0.1307	0.0230	5.68	0.0000
$ECT_{inf,\ t-1}$	-0.3250	0.0684	-4.75	0.0000
$R^2 = 0.63$ $\sigma = 1.72\%$				

Misspecification test		*value*	*p-value*
AR 1-4 test	F(4, 36)	0.3694	0.8288
ARCH 1-4 test	F(4, 32)	0.9778	0.4334
Normality test	$\chi^2(2)$	0.0768	0.9623
Hetero test	F(8, 31)	0.3469	0.9401
Hetero-X test	F(14, 25)	0.3922	0.9643
RESET test	F(1,39)	1.2840	0.2641
Instability test	Var (Joint)	0.0823	(0.6953)

Source: Author's calculations

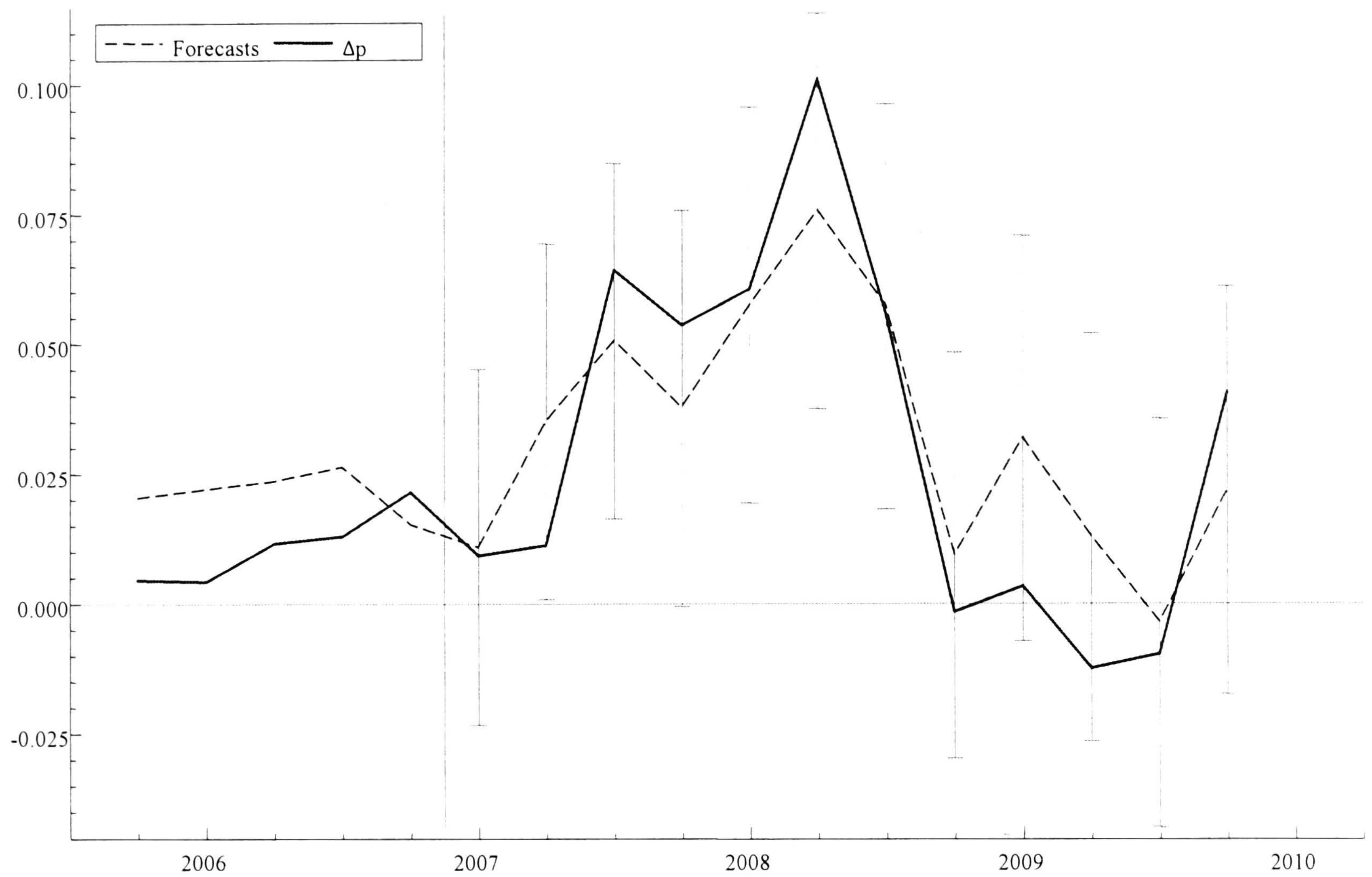
Figure 3.2 Inflation model recursive coefficients with ±2SE, one-step residuals, and constancy tests



Source: Author's calculations

Figure 3.3 plots actual and forecast values of inflation. In order to check the forecasting performance, we re-estimated the inflation model through 2006:4 and forecast inflation for 2007-2009 using actual values of the right-hand side variables. The forecast values track the recent surge in inflation reasonably accurately.

Figure 3.3 Inflation model actual and forecast values with ±2SE



Source: Author's calculations

3.5 Estimation results of the Bayesian approach

Money demand and price markup in this section are estimated by the Bayesian cointegration approach suggested by Villani (2005) and Warne (2006). A number of Bayesian approaches to cointegration are found in the literature, including: Kleibergen and van Dijk (1994); Bauwens and Lubrano (1996); Geweke (1996); Bauwens and Giot (1998); Kleibergen and Paap (2002); Strachen (2003); Strachen and Inder (2004); Villani (2005); and Warne (2006). Villani (2005) suggests a prior for cointegration space rather than a prior for the exactly identified cointegration relations. The paper by Warne (2006) extends the Villani cointegration procedure in two important dimensions:

first, it allows for proper prior distribution of the short-run parameters on lagged endogenous variables; second, an analytical expression of the posterior mode is derived.

We use the reference prior developed by Villani (2005) for the prior distribution. The joint prior distribution is as follows:

$$p(\alpha,\beta,\Phi,\Gamma,\Omega,r)=p(\alpha,\beta,\Phi,\Gamma,\Omega\,|\,r)p(r), \tag{5.1}$$

where $p(r)=1/(p+1)$ for all $r\in\{0,1,...,p\}$. The prior density of $(\alpha,\beta,\Phi,\Gamma,\Omega)$ conditional on the cointegration rank, r, is given by

$$p(\alpha,\beta,\Phi,\Gamma,\Omega|r)=c_r|\Omega|^{-(p+q+r+1)/2}\exp\left(-\frac{1}{2}tr\left[\Omega^{-1}\left(A+\left(1/\lambda_\alpha^2\right)\alpha\beta'\beta\alpha'\right)\right]\right)p(\Gamma|\Omega), \tag{5.2}$$

where $\lambda_\alpha>0$, $q\geq p$, and A, a $p\times p$ positive definite matrix, are three hyperparameters, and c_r is the normalizing constant. We will set $\lambda_\alpha=0.7$, $q=p+2$, and $A=\lambda_A I_p$ where $\lambda_A=0.2$.

Under informative prior of $\Gamma|\Omega$, we follow Warne (2006) and assume that it is matricvariate normal:

$$p(\Gamma|\Omega)=(2\pi)^{-p^2(k-1)/2}|\Sigma_\Gamma|^{-p/2}|\Omega|^{-p(k-1)/2}\exp\left(-\frac{1}{2}tr\left[\Omega^{-1}\Gamma\Sigma^{-1}\Gamma'\right]\right). \tag{5.3}$$

The $p(k-1)\times p(k-1)$ matrix Σ_Γ is positive definite and block diagonal with $(k-1)$ blocks consisting of the $p\times p$ matrices

$$\Sigma_\Gamma=\frac{\lambda_b^2}{i^{2\lambda_l}}I_p,\quad i=1,...,k-1. \tag{5.4}$$

The hyperparameter $\lambda_b>0$ adjusts the overall tightness around zero, while $\lambda_l>0$ measures the lag order shrinkage. In this paper $\lambda_b=1.5$ and $\lambda_l=1$.

With $\theta = (\alpha, \beta, \Phi, \Gamma, \Omega)$, the full conditional posteriors of the five groups of parameters conditional on the rank are given in Warne (2006, Proposition 1). In practice we do not use the full conditional distribution. Instead, we estimate the marginal posterior of α conditional on β and the rank, *r*, and of Ψ[6] conditional on α and the rank, *r*, because it saves computation time for the Gibbs sampler. The marginal conditional posteriors are also provided in Warne (2006, Proposition 2).

Using the Bayesian rule, we can estimate the posterior probability of the cointegration rank:

$$p(r|\mathfrak{D}) = \frac{p(\mathfrak{D}|r)\,p(r)}{\sum_{i=0}^{p} p(\mathfrak{D}|i)\,p(i)} \tag{5.5}$$

The marginal likelihoods, $p(\mathfrak{D}|r)$, for $r = 0$ and $r = p$ have analytical expression, while for $r \in \{1, \ldots, p-1\}$ they do not have closed form expression. Villani (2005) advocates the marginal likelihood identity approach to compute $p(\mathfrak{D}|r)$. The marginal likelihood identity can be written as

$$p(\mathfrak{D}|r) = \frac{p(\mathfrak{D}|\alpha, \Psi, r)\,p(\alpha, \Psi|r)}{p(\Psi|\alpha, \mathfrak{D}, r)\,p(\alpha|\mathfrak{D}, r)}. \tag{5.6}$$

Chib (1995) suggests that the point $(\tilde{\alpha}, \tilde{\Psi})$ should have high posterior density, such as the mode or the median. The posterior mode estimators of the parameters of the posterior $p(\alpha, \Psi|\mathfrak{D}, r)$ are given in Warne (2006, Proposition 4.)

A. Long-run relationship

Money demand

---

[6] $\Psi = \bar{c}_{\perp}'\beta(\bar{c}'\beta)^{-1}$,, c is a known $p \times r$ matrix of rank r with $\bar{c} = c(c'c)^{-1}$.

Under the informative prior of $\Gamma \mid \Omega$, the lag order selection is based on Warne (2006). Almost all the posterior probability mass is given to $k = 1$, the lag order selected by Schwarz and Hannan-Quinn information criteria.

The analysis of the cointegration rank conditional on $k = 1$ is presented in Table 3.6. After 1000 burn-in draws, an additional 5000 draws were simulated using Gibbs sampler for estimating posterior draws of β. The convergence of the sampler has been tested by looking at the recursive posterior median point estimates of the parameters. From Table 3.6 we find that under informative and non-informative prior of $\Gamma \mid \Omega$ there is only one cointegration relation among the real money, output, and interest rate.

Table 3.6 Posterior Cointegration Rank Probabilities for money demand

***Bayesian Informative* $\Gamma|\Omega$**

r	P(r|D)	ln[P(D|r)]	P(r)	st.err. ln[P(D|r)]
0	0.0039	-411.7744	0.25	-
1	0.9910	-406.2359	0.25	0.0080
2	0.0051	-411.5112	0.25	0.0243
3	0.0000	-420.0522	0.25	-

***Bayesian Non-Informative* $\Gamma|\Omega$**

r	P(r|D)	ln[P(D|r)]	P(r)	st.err. ln[P(D|r)]
0	0.0038	-411.7744	0.25	-
1	0.9912	-406.2066	0.25	0.0199
2	0.0050	-411.4884	0.25	0.0238
3	0.0000	-420.0522	0.25	-

Source: Author's calculations

Table 3.7 presents point estimates and 95 percent confidence bands for the money demand equation. All these estimates are similar to the maximum likelihood estimates provided in section 3.4. In particular, posterior modes are very close to classical estimates under both priors.

Table 3.7 Bayesian point estimates of long-run money demand

***Bayesian Informative* $\Gamma|\Omega$**

	real money	output	interest rate
Posterior mode	1	-1.0049	6.5403
Posterior median	1	-1.0435	5.0385
Posterior mean	1	-1.0340	5.0293
95 percent confidence bands	-	[-1.1394, -0.8318]	[4.1279, 6.2426]

***Bayesian Non-Informative* $\Gamma|\Omega$**

	real money	output	interest rate
Posterior mode	1	-1.0049	6.5403
Posterior median	1	-0.8533	6.1091
Posterior mean	1	-0.8870	6.0508
95 percent confidence bands	-	[-1.1530, -0.7652]	[4.7409, 7.7554]

Source: Author's calculations

Markup

Under the informative prior of $\Gamma \mid \Omega$, approximately all the posterior probability mass is given to $k=1$. Conditional on this lag order, $k=1$, the analysis of the cointegration rank of price markup is shown in Table 3.8. After 1000 burn-in draws, an additional 5000 draws were simulated using the Gibbs sampler for estimating posterior draws of β. From Table 3.8 we find that there is only one cointegrating vector of price markup.

Table 3.8 Posterior Cointegration Rank Probabilities for markup

***Bayesian Informative* $\Gamma|\Omega$**

r	P(r|D)	ln[P(D|r)]	P(r)	st.err. ln[P(D|r)]
0	0.1243	-565.2951	0.2	-
1	0.8753	-563.343	0.2	0.0324
2	0.0004	-571.0454	0.2	0.1087
3	0	-579.6778	0.2	0.0273
4	0	-590.738	0.2	-

***Bayesian Non-Informative* $\Gamma|\Omega$**

r	P(r|D)	ln[P(D|r)]	P(r)	st.err. ln[P(D|r)]
0	0.115	-565.2951	0.2	-
1	0.8847	-563.2549	0.2	0.0384
2	0.0003	-571.1048	0.2	0.1047
3	0	-579.6964	0.2	0.0274
4	0	-590.738	0.2	-

Source: Author's calculations

Table 3.9 presents point estimates and 95 percent confidence bands for the markup equation. All of these estimates are similar to the maximum likelihood estimates provided in section 4. Under non-informative prior of $\Gamma \mid \Omega$, the posterior median and mean of import price are smaller than that of Johansen's maximum likelihood estimates.

Table 3.9 Bayesian point estimates of long-run price

***Bayesian Informative* $\Gamma|\Omega$**

	price	ulc	ip	pet
Posterior mode	1	-0.3325	-0.4327	-0.2148
Posterior median	1	-0.3321	-0.4247	-0.2150
Posterior mean	1	-0.3293	-0.4190	-0.2192
95 percent confidence bands	-	[-0.4169,-0.2297]	[-0.5229,-0.2980]	[-0.3250,-0.1330]

***Bayesian Non-Informative* $\Gamma|\Omega$**

	price	ulc	ip	pet
Posterior mode	1	-0.3325	-0.4327	-0.2148
Posterior median	1	-0.3316	-0.2416	-0.2523
Posterior mean	1	-0.3277	-0.2439	-0.2550
95 percent confidence bands	-	[-0.4490,-0.1939]	[-0.3810,-0.1312]	[-0.3937,-0.1236]

Source: Author's calculations

In sum, we conclude that Bayesian estimates of cointegration for money demand and price markup equations are similar to the maximum likelihood estimates provided in section 3.4.

B. Short-run inflation dynamics

Bayesian model averaging has become an important tool in empirical research with large numbers of regressors and relatively limited numbers of observations. Here we investigate short run inflation dynamics using the Bayesian model averaging approach developed by Fernandez, Ley and Steel (2001), who propose a benchmark prior distribution that works for the general condition which includes substantive prior information into the analysis.

Following Fernandez et al. (2001) we denote by M_j the model with regressors stacked into Z_j

$$y = \alpha i_n + Z_j \beta_j + \sigma\varepsilon \qquad (5.7)$$

where $\beta_j \in \Re^{k_j} (0 \le k_j \le k)$ stacks the regression coefficients, $\sigma \in \Re_+$ is a scale parameter and ε follows a normal distribution, $N(0, I_n)$. A prior distribution for the model, M_j follows as

$$p(\alpha, \sigma) \propto \sigma^{-1} \tag{5.8}$$

and

$$p(\beta_j \mid \alpha, \sigma, M_j) = f_N^{k_j}(\beta_j \mid 0, \sigma(gZ_j'Z_j)^{-1}) \tag{5.9}$$

where $f_N^{k_j}$ denotes k_j-dimensional Normal distribution. Fernandez et al. (2001) investigate nine possible choices for g in equation (5.9) using a Monte Carlo simulation and propose a simple rule. According to their study, the simple rule, $g = 1/\max\{n, k^2\}$, works for the simulation. We also specify a prior distribution over the space of 2^k possible models as follows:

$$P(M_j) = p_j = 2^{-k}, \quad j = 1, \ldots, 2^k, \quad p_j > 0, \quad \sum_{j=1}^{2^k} p_j = 1 \tag{5.10}$$

With the Bayesian approach, it is simple to estimate the posterior distribution of any quantity of interest, Δ:

$$P_{\Delta|y} = \sum_{j=1}^{2^k} P_{\Delta|y, M_j} P(M_j \mid y) \tag{5.11}$$

This Bayesian Model Averaging (BMA) equation follows from the rules of probability theory. Posterior model probabilities are given by:

$$P(M_j \mid y) = \frac{l_y(M_j) p_j}{\sum_{h=1}^{2^k} l_y(M_h) p_h} \tag{5.12}$$

where $l_y(M_j)$, the marginal likelihood of model M_j is obtained as

$$l_y(M_j) \propto \left(\frac{g}{g+1}\right)^{k_j/2} \left(\frac{1}{g+1} y' M_{x_j} y + \frac{g}{g+1} (y - \bar{y} i_n)'(y - \bar{y} i_n)\right)^{-(n-1)/2} \tag{5.13}$$

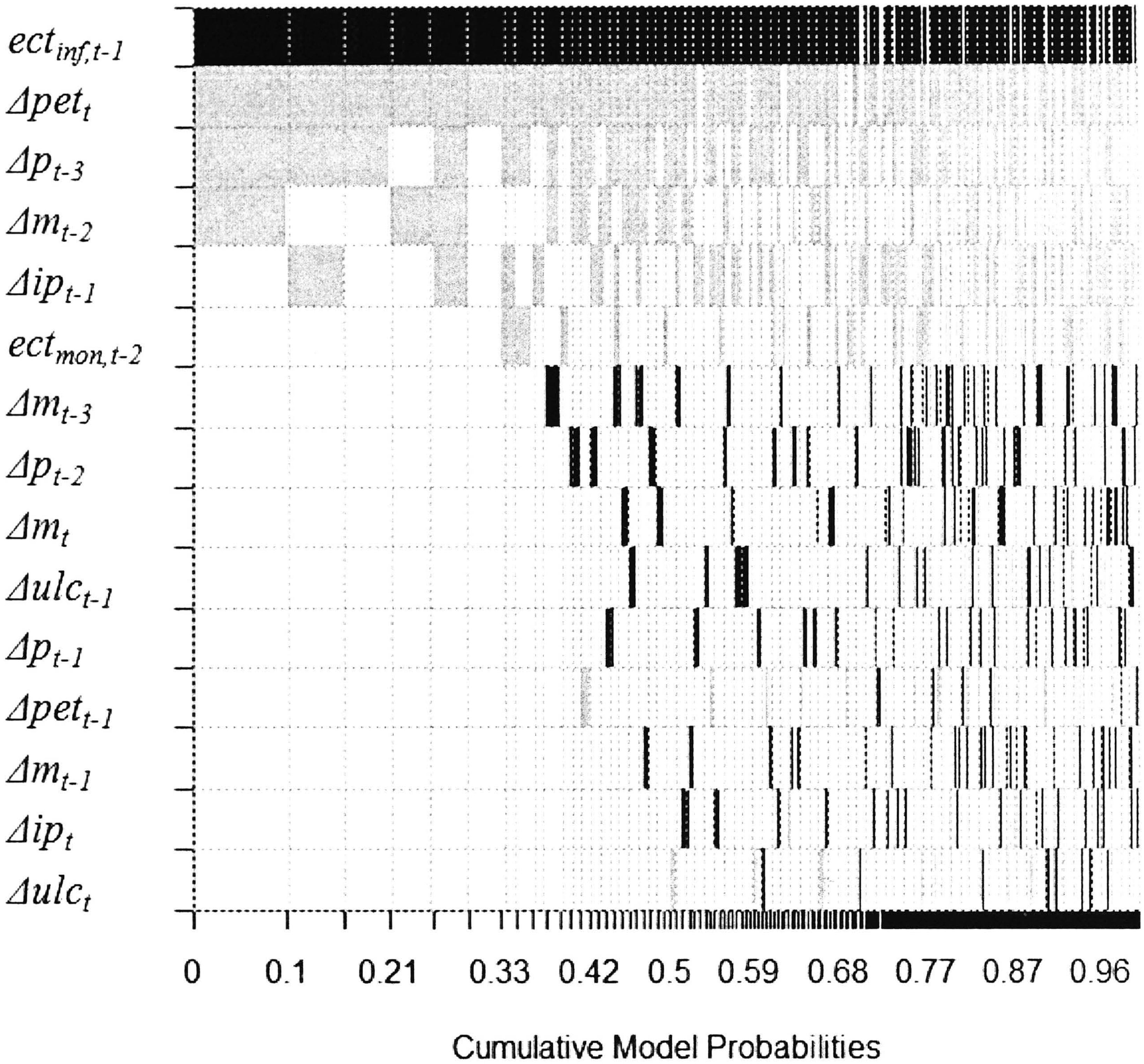
where $i_n' Z = 0$, $X_j = (i_n : Z_j)$, $\bar{y} = i_n' y / n$ and $M_{X_j} = I_n - X_j (X_j' X_j)^{-1} X_j'$.

In order to substantially reduce the computational effort, we will use the MCMC model composition (MC^3) method following Fernandez et al. (2001). This Metropolis algorithm works as follows. Given that the chain is currently at model M_s, a new model M_j is proposed randomly through a Uniform distribution on the space containing M_s, and all models with either one regressor more or one regressor less than M_s. The chain moves to M_j with probability $p = \min\{1, [l_y(M_j) p_j] / [l_y(M_s) p_s]\}$ and remains at M_s with probability $1 - p$.

Since $n = 44 < k^2 = 225$ we will use $g = 1/k^2$ in the prior in (5.9). Our results are based on 50000 drawings after a burn-in of 10000 drawings. The correlation coefficient between posterior model probabilities based on empirical visit frequencies and analytical formulae based (5.13) is 0.996, suggesting the chain has reached its equilibrium distribution.

Of all possible models of $2^k = 32768$, 16228 models are visited and the best model accounts for a posterior probability of 10%. The mass is spread out: the best 100 models cover 80% of the posterior model probability.

Figure 3.4 Model Inclusion based on best models



Source: Author's calculations

Posterior probability of inclusion of variables and best models are shown in Figure 3.4. In the figure, the light grey colour corresponds to a positive coefficient, black to a negative coefficient, and white to non-inclusion. On the horizontal axis are shown the best models, scaled by their posterior model probabilities. For instance, the best model selected by Bayesian model averaging is the same as the specific model estimated in section 3.4. The next best model selects inflation inertia, import price lagged one quarter, petroleum price, and an error correction term for price markup.

Figure 3.4 also shows that the change of petroleum price and the price markup (error correction term) are the most important determinants for inflation dynamics with posterior probabilities of inclusion at 93 % and 96 %, respectively. The other important determinants are inflation inertia lagged three quarters (69 %), the change in import price lagged one quarter (42 %), the change in M1 lagged two quarters (46 %), and excess supply of money lagged two quarters (20 %).

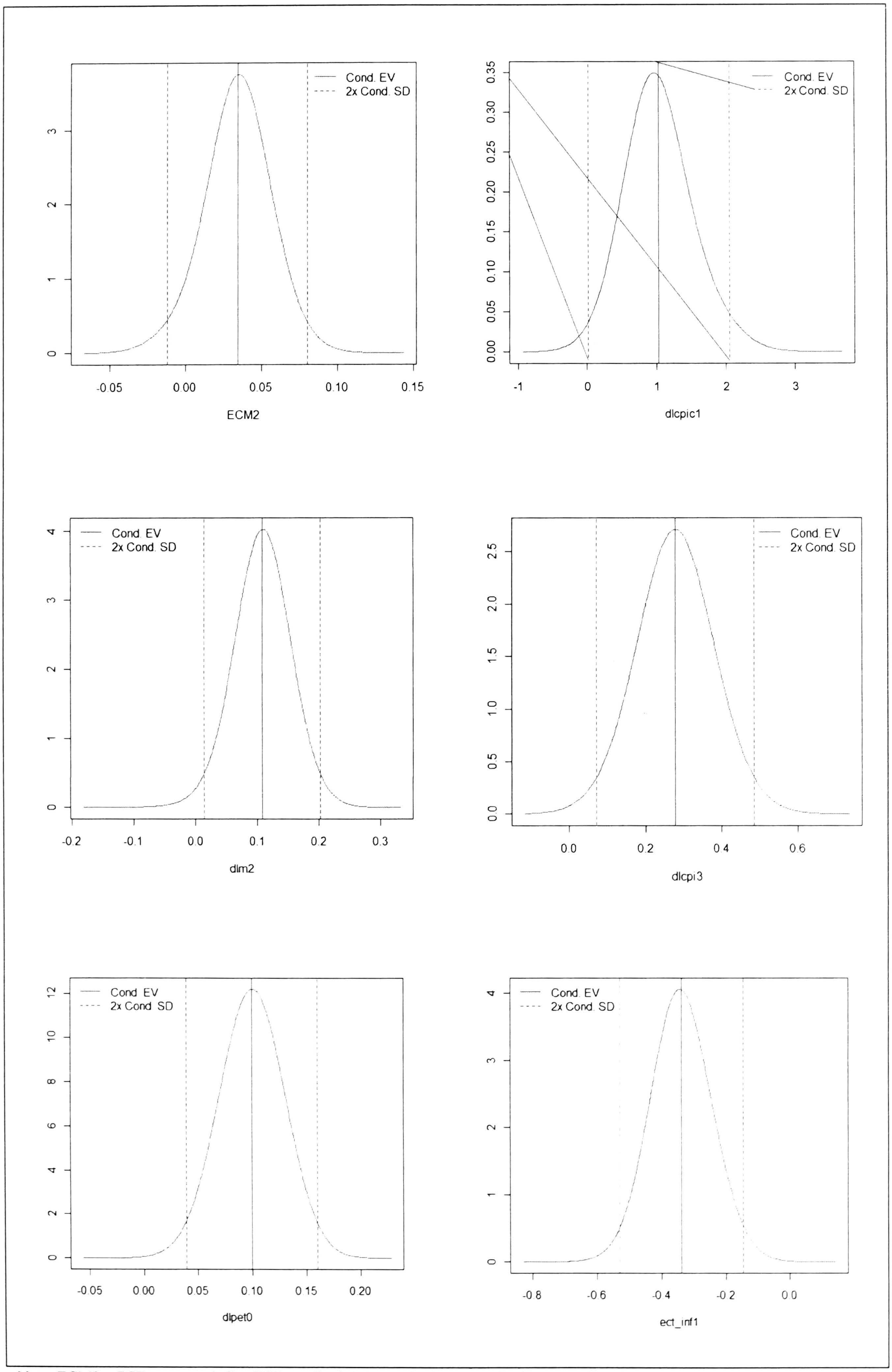
Table 3.10 Bayesian Model Averaging: Conditional Posterior Moments

		PIP	Post mean	Post SD	Cond.Post.Sign
1	Δp_{t-1}	0.0644	-0.0719	0.1115	0.0103
2	Δp_{t-2}	0.0934	-0.1200	0.1107	0.0000
3	Δp_{t-3}	0.6963	0.2782	0.1031	1.0000
4	Δulc_t	0.0571	0.0085	0.0430	0.6850
5	Δulc_{t-1}	0.0681	-0.0210	0.0539	0.1141
6	Δip_t	0.0575	-0.1608	0.4602	0.1953
7	Δip_{t-1}	0.4096	1.0159	0.4987	1.0000
8	Δpet_t	0.9369	0.1001	0.0302	1.0000
9	Δpet_{t-1}	0.0616	0.0153	0.0306	0.8615
10	$\Delta m1_t$	0.0708	-0.0340	0.0442	0.0043
11	$\Delta m1_{t-1}$	0.0611	-0.0249	0.0467	0.0231
12	$\Delta m1_{t-2}$	0.4645	0.1089	0.0466	0.9989
13	$\Delta m1_{t-3}$	0.1019	-0.0517	0.0470	0.0025
14	$ECT_{inf,t-1}$	0.9678	-0.3372	0.0953	0.0000
15	$ECT_{mon,t-2}$	0.1869	0.0336	0.0226	1.0000

Note: PIP column represents posterior inclusion probabilities.
Source: Author's calculations

Bayesian model averaging estimates of short run inflation dynamics are provided in Table 3.10. Conditional estimates are the posterior distribution moments of the regression coefficient, given that the corresponding variable is included in the regression. We also illustrate the marginal posterior distribution of some regression coefficients with expected values and two times standard deviation bounds in Figure 3.5. From the figure, it is clear that the coefficient of the error correction term for money is above zero. Thus, excess money seems to determine inflation in the long run but the adjustment to disequilibria is slow. Further, the change in import price lagged one quarter seems to have a short run effect on inflation. In brief, Bayesian model averaging identifies two additional determinants of excess money and change in import prices compared with the classical analysis (see section 3.4).

Figure 3.5 Posterior densities of selected coefficients



Note: ECM2 - $ECT_{mon,t-2}$; ect_infl - $ECT_{inf,t-1}$;dlcpic1-Δip_{t-1}; dlm2 - Δm_{t-2}; dlcpi3 - Δp_{t-3}; dlpet0 - Δpet_t
Source: Author's calculations

Since the estimated coefficient on lagged inflation reflects to some degree the formation of inflation expectations, we have an economic reason to suspect it will change depending on inflation expectation. We model this possible nonlinearity of inflation persistence by the Markov switching model (see the appendix).

3.6 Conclusions

In order to better understand inflation dynamics and the policy that a central bank implements, this paper has developed a single-equation error correction model for inflation in Mongolia. The resulting model is interpretable, parsimonious and empirically stable.

The main findings of the paper are summarized as follows. First, the main determinant of inflation is the markup, capturing impacts from labour costs, petroleum prices, import prices, and the exchange rate in the long run. For the short run, inflation inertia, petroleum prices and nominal money (M1) growth describe inflation dynamics well. Second, an excess money supply in the money market seems to cause inflation in the long run if the model uncertainty and nonlinearity are considered but the adjustment to disequilibria is slow. Third, the sustained increases in wages, coupled with the petroleum price shock, explain the high and volatile rate of inflation experienced in recent years. In contrast, the appreciation of the exchange rate due to commodity price hikes has contributed to containing inflation. Finally, we model the possible nonlinearity of inflation persistence by the Markov switching model and find two inflationary regimes, which are characterized by a high and a low degree of inflation persistence.

This paper draws some policy implications for Mongolia. First, inflation could be reduced by keeping wages consistent with the level suggested by productivity. Second, the central bank should track narrow money (M1) rather than broad money (M2). Third, policy makers need to consider inflationary regimes when making decisions about monetary policy.

3.7 Appendix

Inflation persistence and nonlinearity

In this appendix we model the possible nonlinearity of the inflation persistence using the Markov switching model. Use of the Markov switching model has proliferated since Hamilton's (1989) seminal paper on business cycle dating. A nice overview of regime Markov-switching model can be found in Hamilton (1994). The monographs by Kim and Nelson (1999) and Frühwirth-Schnatter (2006) provide detailed techniques of the Bayesian approach of Markov switching models. We estimate following the best model selected by a general-to-specific approach and Bayesian model averaging.

$$\pi_t = \beta_0 + \beta_{1,S_t}\pi_{t-3} + \beta_2 \Delta pet_t + \beta_3 \Delta m_{t-2} + \beta_4 ect_{\inf,t-1} + \varepsilon_t \qquad (6.1)$$

with $\varepsilon_t \sim N(0, \sigma^2_{S_t})$. Here π_t denotes inflation. Inflation inertia or persistence coefficient β_{1,S_t} and variance $\sigma^2_{S_t}$ will depend on random discrete variable $s_t = 1,2$.

Classical approach

We do not observe s_t directly and make an inference using two probabilities

$$\xi_{jt} = \Pr(s_t = j \mid \Omega_t; \theta) \qquad (6.2)$$

Here Ω_t denotes a set of observations and θ is a vector of population parameters. Inference is performed iteratively using Hamilton's filter. First, we calculate the conditional density of the t th observation from

$$f(\pi_t \mid \Omega_{t-1}; \theta) = \sum_{i=1}^{2}\sum_{j=1}^{2} p_{ij}\xi_{i,t-1}\eta_{jt} \qquad (6.3)$$

Here p_{ij} are transition probabilities and η_{jt} the densities under two regimes . Then the desired probabilistic inference is

$$\xi_{jt} = \frac{\sum_{i=1}^{2} p_{ij}\xi_{i,t-1}\eta_{jt}}{f(\pi_t \mid \Omega_{t-1};\theta)} \tag{6.4}$$

By the product of this filter, we will obtain a sample conditional log likelihood of the observed data

$$L = \sum_{t=1}^{T} \log f(\pi_t \mid \Omega_{t-1};\theta) \tag{6.5}$$

for the specified value of θ. An estimate of the value of θ is calculated by maximizing log likelihood by numerical optimization. Maximum likelihood estimates of parameters are reported in Table 3.11a.

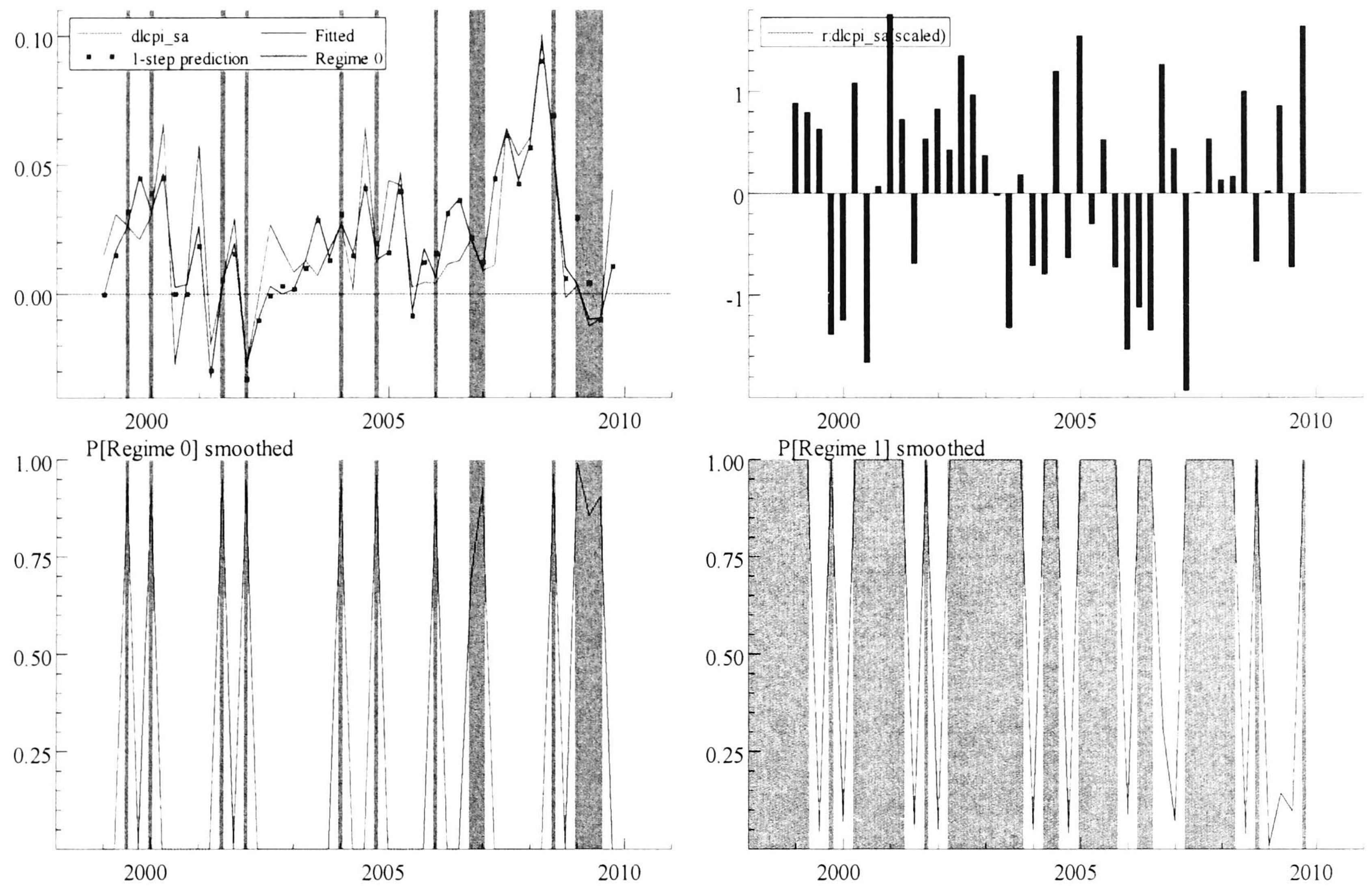
Table 3.11a Markov switching model

	Coefficient	Standard error	t-value	t-prob
constant	0.0022	0.0002	10.70	0.00
$\pi_{t-3,1}$	0.0780	0.0040	19.30	0.00
$\pi_{t-3,2}$	0.4581	0.1058	4.33	0.00
Δpet_t	0.1199	0.0011	108.00	0.00
Δm_{t-2}	0.1071	0.0022	48.80	0.00
$ect_{\inf,t-1}$	-0.4145	0.0043	-96.60	0.00
σ_1	0.0004	0.0001	5.48	0.00
σ_2	0.0179	0.0023	7.96	0.00
p_{11}	0.1804	0.1383	1.30	0.20
p_{21}	0.3074	0.0884	3.48	0.00

Note: Linearity LR-test $\chi^2(4)$ = 38.096 [0.0000] approximate upperbound: [0.0000]
Source: Author's calculations

In the regime represented by $s_t = 1$, the inflation persistence coefficient is $\beta_{1,s_t=1} = 0.078$, while when $s_t = 2$ the inflation persistence coefficient is $\beta_{1,s_t=2} = 0.458$. Regime 2 is highly persistent. The probability that high inflation persistence will be followed by another high inflation persistence is $p_{22} = 1 - p_{21} = 0.69$ so that this regime will persist on average for $1/(1-p_{22}) = 3$ quarters. The probability that low inflation inertia will be followed by low inflation inertia is $p_{11} = 0.18$, which episodes will persist for $1/(1-p_{11}) = 1$ quarter.

Figure 3.6a Markov switching model



Source: Author's calculations

A probabilistic inference in the form of (6.2) can be calculated for each date t in the sample. Figure 3.6a depicts the smoothed probabilities of inflation persistence regime. Given parameter estimates of the model, the smoothed probabilistic inference uses all the information in the sample up to a later date *T*. An efficient algorithm for smoothed probabilities, $\xi_{t|T}$ was developed by Kim (1994).

It is natural to test whether the non-linearity adds anything to the linear, constant-parameter model. Such a test for linearity fails to satisfy the usual regularity condition because parameters are not identified under the null hypothesis. As a consequence, the likelihood ratio test does not have the standard χ^2 limiting distribution.

We report a test for linearity, which is based on the likelihood-ratio statistic between the derived linear model and the estimated model. The first p-value is based on the usual χ^2

distribution. Also reported is the approximate upperbound for the significance level of the LR statistic as derived by Davies (1977). Both tests reject linearity of the model[7].

It is also interesting to test whether excess money significantly affects inflation once we allow for nonlinearity of inflation persistence. Table 3.12a presents the regression results. The Markov switching model suggests that inflation responds significantly to excess money.

Table 3.12a Markov switching model with excess money

	Coefficient	Standard error	t-value	t-prob
constant	0.0166	0.0006	26.40	0.00
$\pi_{t-3,1}$	0.0711	0.0100	7.12	0.00
$\pi_{t-3,2}$	0.2283	0.1220	1.87	0.07
Δpet_t	0.0889	0.0031	28.90	0.00
Δm_{t-2}	0.0889	0.0043	20.80	0.00
$ect_{\text{inf},t-1}$	-0.3345	0.0096	-34.70	0.00
$ect_{\text{mon},t-2}$	0.0476	0.0026	18.00	0.00
σ_1	0.0009	0.0002	4.28	0.00
σ_2	0.0206	0.0027	7.66	0.00
p_{11}	0.4528	0.1580	2.87	0.01
p_{21}	0.2327	0.1026	2.27	0.03

Source : Author's calculations

Bayesian approach

In the Bayesian analysis, both the parameters θ of the models and the Markov-switching variable $s=(s_1,s_2,\ldots,s_T)'$ are treated as random variables. Albert and Chib (1993) have developed an algorithm of the Bayesian approach of Markov-switching models using the simulation tool of Gibbs-sampling. This is estimated by sequentially generating a realization of $s^{(k)}$ from the distribution of $s\mid\theta^{(k-1)},\Omega_T$ followed by a realization $\theta^{(k)}$ from the distribution of $\theta\mid s^{(k)},\Omega_T$.

The parameter estimates are summarized in Table 3.13a. We assume no prior correlation among the regression parameters. We compare the Markov switching model

[7] One can use other testing methods of Garcia (1998) and Breunig and Pagan (2004).

with the linear model of inflation using the marginal likelihoods. Marginal likelihoods are estimated from the Gibbs sampling output of a random permutation sampler (6000 simulations after burn-in iteration of 1000) using the optimal bridge sampling estimators. The log of marginal likelihood of the Markov switching model is higher than that of the linear model (99.04 > 96.45). Thus, we reject the null hypothesis that the model is linear against the alternative that it is nonlinear.

Table 3.13a Bayesian Markov switching model

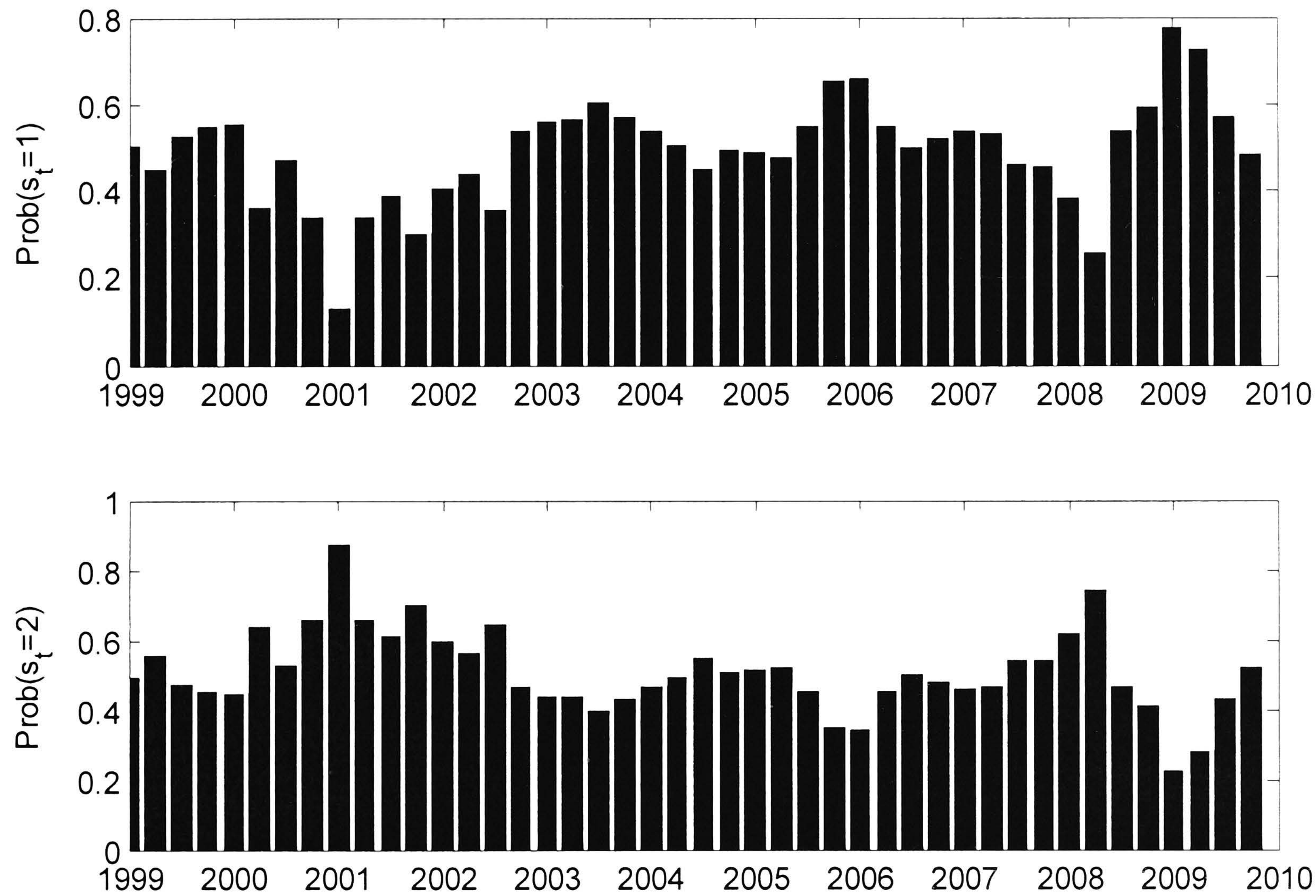
	Prior		Posterior	
	mean	standard dev	mean	standard dev
Constant	0.0000	2.0000	0.0059	0.0039
$\pi_{t-3,1}$	0.0000	0.1000	0.0764	0.0865
$\pi_{t-3,2}$	0.4000	0.1000	0.3690	0.0887
Δpet_t	0.0000	2.0000	0.1134	0.0269
Δm_{t-2}	0.0000	2.0000	0.1129	0.0457
$\Delta ect_{inf,t-1}$	0.0000	2.0000	-0.3389	0.0732
σ_1	0.0004	0.0006	0.0003	0.0001
σ_2	0.0004	0.0006	0.0003	0.0001
p_{11}	0.7000	0.1382	0.6932	0.1357
p_{22}	0.7000	0.1382	0.6960	0.1385

Note: the prior on β follows normal distribution, the prior on σ2 inverse gamma distribution and the prior on transition probability matrix, p beta distribution.
Source: Author' calculations

Compared with the maximum likelihood estimates, Bayesian estimates are not much different except for the transition probability estimates for regime 1. The probability estimate for the low inflation inertia state is almost equal to that for the high inflation persistence state, meaning the duration of the two states is approximately three quarters.

In Figure 3.7a, the posterior probability of the state of the inflation persistence, $\Pr(s_t = j \mid \Omega_t)$, is graphed. Compared with classical inferences for smoothed probabilities, Bayesian inferences are not conditional on the parameter estimates. In the classical framework, inference on the Markov switching models consists of first estimating a model's unknown coefficients, then making inferences on the unobserved state variable. Therefore the smoothed state probabilities in Figure 3.7a are somewhat different from those in Figure 3.6a.

Figure 3.7a Smoothed state probabilities



Source: Author's calculations

In sum, we find two inflationary regimes, characterized by the degree of inflation persistence.

Table 3.14a ADF test statistics

	Level			First differences		
	test stat	p-value	lag	test stat	p -value	lag
m1	-0.3107	0.9156	0	-6.0994	0.0000	0
m1-p	-0.9212	0.7732	0	-6.7005	0.0000	0
P	1.6760	0.9995	5	-3.6613	0.0083	3
Y	0.6215	0.9888	3	-7.7885	0.0000	2
*i**	-0.4531	-1.9477	0	-5.8983	-1.9477	0
ulc	-0.3828	0.9038	0	-7.3459	0.0000	0
ip	0.3819	0.9801	1	-3.6761	0.0076	0
pet	-0.9277	0.7707	0	-7.5091	0.0000	0

Note: * Dickey Fuller GLS test and 5 percent critical values are reported.

Source: Author's calculation

Table 3.15a VAR lag order selection

Money demand (1997:4 2009:4)				
Lag	LogL	LR	SC	HQ
0	128.7035	NA	-5.5921	-5.6687
1	263.6975	245.4437	**-10.9542**	**-11.2603**
2	273.5228	16.5244	-10.6268	-11.1625
3	280.3116	10.4918	-10.1613	-10.9267
4	299.4077	**26.9081**	-10.2553	-11.2502
Price (1998:1 2009:4)				
Lag	LogL	LR	SC	HQ
0	197.1532	NA	-8.6175	-8.7195
1	374.4187	**314.2435**	**-15.2990**	**-15.8092**
2	389.1124	23.3762	-14.5908	-15.5092
3	400.8982	16.6073	-13.7504	-15.0770
4	421.0473	24.7285	-13.2902	-15.0250

Note: LR: sequential modified LR test statistic (each test at 5% level); SC: Schwarz information criterion; HQ: Hannan-Quinn information criterion

Source: Author's calculations

Table 3.16a Misspecification tests

	Money demand		
Serial correlation test	*type*	*value*	*p-value*
LM1	F(9, 36)	1.1490	0.3558
LM4	F(9, 33)	1.5077	0.1862
Normality test	*type*	*value*	*p-value*
Wald type test	F(6, 38)	0.7631	0.6034
Shenton-Bowman/ Doornik-Hansen test	F(6, 38)	1.3783	0.2483
ARCH test	*type*	*value*	*p-value*
Multivariate	F(36, 40)	1.9555	0.0202
	Price		
Serial correlation test	*type*	*value*	*p-value*
LM1	F(16, 37)	1.2378	0.2871
LM4	F(16, 34)	1.3920	0.2036
Normality test	*type*	*value*	*p-value*
Wald type test	F(8, 40)	1.9072	0.0858
Shenton-Bowman/ Doornik-Hansen test	F(8, 40)	1.9871	0.0733
ARCH test	*type*	*value*	*p-value*
Multivariate	F(100, 42)	1.2123	0.2444

Source: Author's calculations

Table 3.17a Cointegration analysis: Nominal money

Johansen test				
Eigenvalue	0.6528	0.2811	0.2240	0.0271
Null hypothesis, H_0	$r = 0$	$r \leq 1$	$r \leq 2$	$r \leq 3$
LR trace	73.4395	26.8891	12.3683	1.2072
Asymptotic p-value	0.0000	0.1044	0.1401	0.2719
Bootstrap p-value	0.0010	0.1692	0.1772	0.3323
LR trace (Bartlett corrected)	63.1805	23.2651	10.3957	0.6249
Asymptotic p-value	0.0010	0.2333	0.2515	0.4292
Bootstrap p-value	0.0010	0.1542	0.1892	0.3604

Standardized cointegration vector beta coefficients

Variables	*m1*	*p*	*y*	*i*
Cointergrating vector , β'	1.0000	0.4366	-2.4144	7.4990

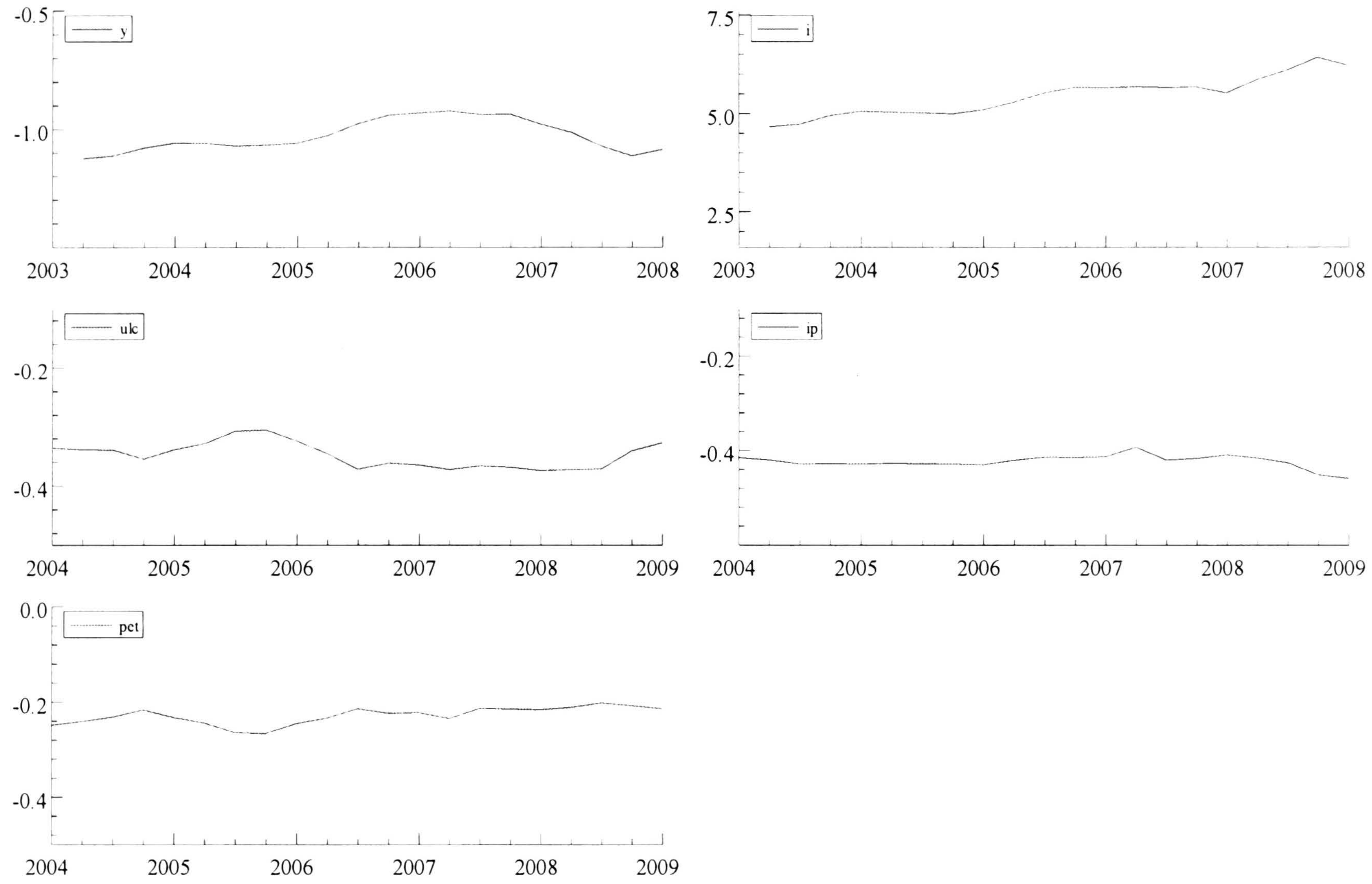
Weak exogeneity test				
$\chi^2(1)$	0.7765	0.8731	3.3951	28.3364
asym. p-value	0.3782	0.3501	0.0654	0.0000

Hypothesis	χ^2	*deg.free*	*p-value*	*boot. p-val*
$H_0 : \beta'=(1\ -1\ -1\ *)$	5.7693	2	0.0559	0.1381

Restricted cointegrating vector	*m1*	*p*	*y*	*i*
Cointegrating vector, β'	1	-1	-1	5.8059
standard error	-	-	-	0.4606

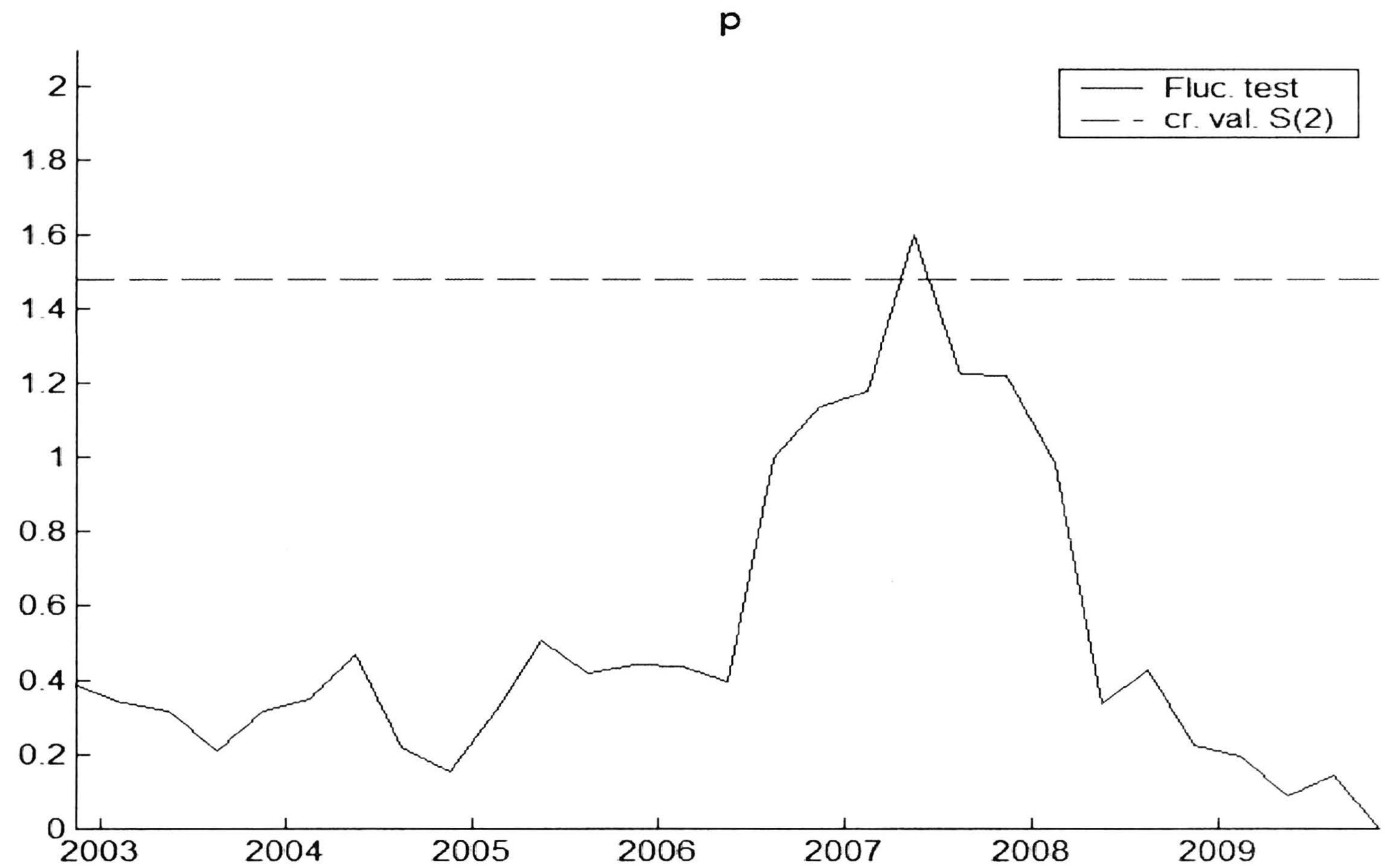
Source: Author's calculations

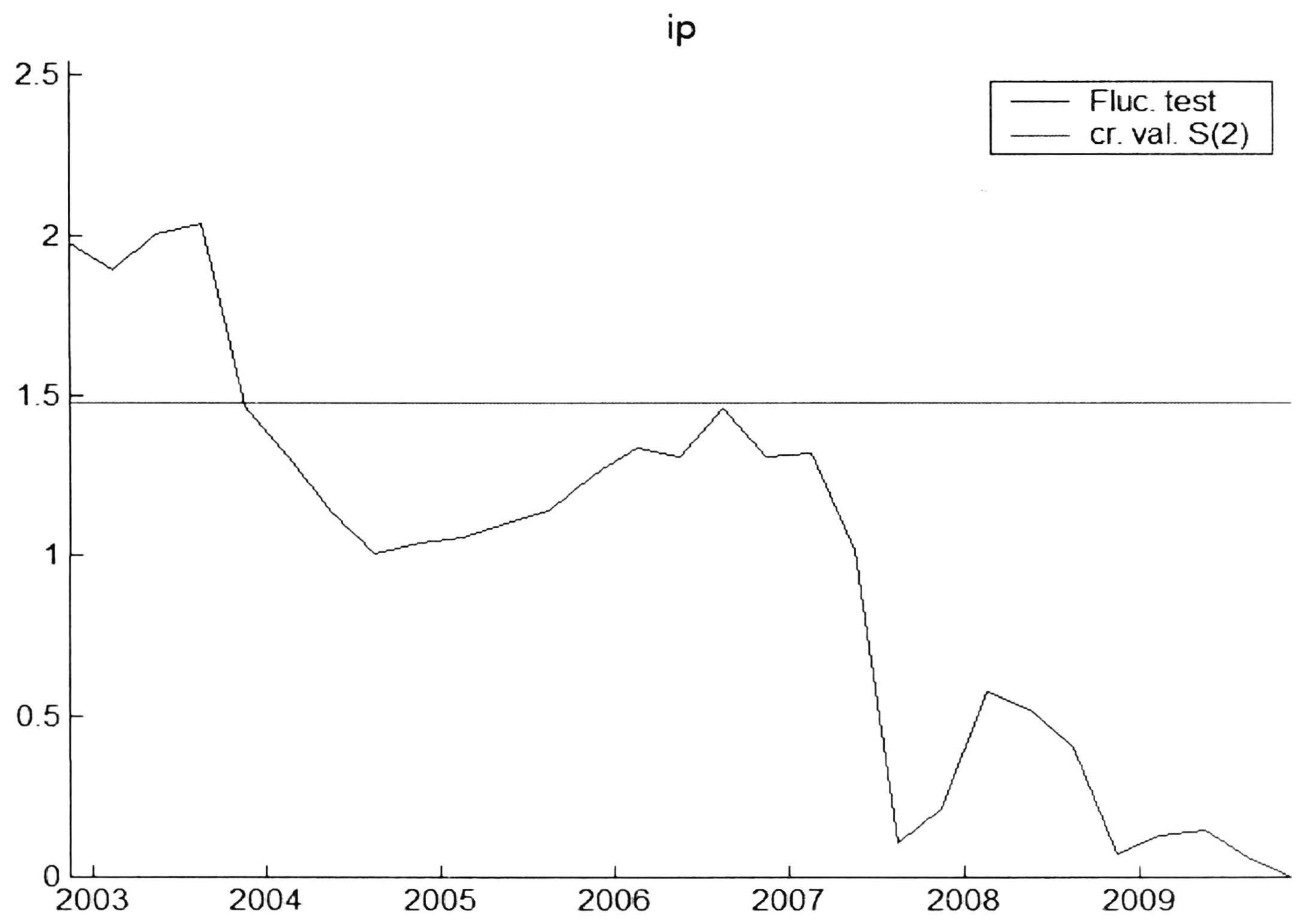
Figure 3.8a Recursively estimated coefficients of money demand and price model



Source: Author's calculations

Figure 3.9a Ploberger-Kramer-Kontrus-tests for p and ip with 95 % critical value





Source: Author's calculations

Chapter 4 MONETARY POLICY REACTION FUNCTION IN MONGOLIA

4.1 Introduction

Nowadays many central banks use Taylor-type rules whereby they respond to output, inflation and exchange rate movements. While there is a large international literature on central bank reaction functions, very little research has been done on how the Bank of Mongolia sets interest rates, although this is important in assessing monetary policy implementation in Mongolia.

Until 1990 Mongolia was a socialist country with a centrally-planned economic system. After the collapse of the Soviet regime, Mongolia moved to a market economy and established a two-tier banking system. The State Bank of Mongolia became the central bank (the Bank of Mongolia) whose main objective was to ensure price stability. This paper estimates the reaction function of the Bank of Mongolia using a Bayesian approach by estimating the New Keynesian dynamic stochastic general equilibrium (DSGE) model of a small open economy developed by Gali and Monacelli (2005), and modified for estimation purposes by Lubik and Schorfheide (2007).

The Bayesian approach has been applied to DSGE models for two different purposes. First, it has been used to facilitate the estimation of a DSGE model with prior information. There are many advantages of using Bayesian methods to estimate DSGE models. From a practical standpoint, the most important is that the inclusion of priors facilitates the identification of model parameters. A problem usually occurs when the posterior distribution is flat over the subspace of parameter values. Moreover, the standard errors of estimates are notoriously difficult to compute and their asymptotic distributions are a poor approximation for a small sample. But combining the likelihood with prior densities often adds enough curvature on the posterior distribution to help numerical maximization. Second, it has been employed to generate forecasts produced by reduced-form vector autoregression (VAR) models. Policymakers make decisions based on forecasts of key variables such as inflation and output. However the forecasting performance of DSGE models is not superior to VAR models due to their restriction. In this regard, the paper uses the DSGE-VAR approach proposed by Del Negro and Schorfheide (2004).

The main finding of this paper is summarized as follows. First, the monetary policy reaction function is forward-looking in terms of the inflation. The expected inflation rule fits the reaction function better than a simple Taylor type rule. Second, the central bank of Mongolia has implemented strong anti-inflationary and exchange rate stabilization policies. Third, there is evidence that the Bank of Mongolia does not respond significantly to output according to the Bayesian posterior odds. Furthermore, the degree of interest rate smoothing is relatively high. The paper consists of seven sections. Section 4.2 introduces the New Keynesian small open economy model, and section 4.3 explains Bayesian estimation of the structural model. The data and priors are described in section 4.4, while empirical results are shown in section 4.5. Finally, section 4.6 demonstrates the forecasting performance of the DSGE-VAR, and section 4.7 concludes the paper.

4.2 New Keynesian small open economy model

Lane (1999) provides an excellent survey of earlier work on optimizing open economy models with nominal rigidities that focus on the transmission of monetary policy shocks. The main contribution in this area is that of Obstfeld and Rogoff (1995, 1996) who develop a two country model where monopolistically competitive firms set prices one period in advance.

Recent papers such as Obstfeld and Rogoff (2002), Benigno and Benigno (2003), Sutherland (2003), Devereux and Engel (2003) and Corsetti and Pesenti (2005) have focused on the implication of two country sticky price open economy models for the design of optimal monetary policy using a welfare approach.

More recent frameworks have adopted the staggered price setting structure of Calvo (1983). The assumption of staggered price and wage setting introduces more realistic dynamics than that of price setting one period in advance. One unsatisfactory feature of price setting one period in advance is its implication that only unanticipated monetary policies have any effect on real output. Gali and Monacelli (2005), Clarida, Gali and Gertler (2001), Kollmann (2002) and Monacelli (2005) develop small open economy models using the Calvo price setting.

Finally, Clarida, Gali and Gertler (2002), Pappa (2004) and Benigno and Benigno (2006) analyse the alternative monetary policy arrangement in a two country framework with a Calvo staggered price setting and with a focus on the gains from cooperation.

The model of this paper is based on a small open economy model, as developed by Gali and Monacelli (2005). However, for estimation purposes it modifies this model as did Lubik and Schorfheide (2007), because a stochastic singularity problem arises if the number of shock terms in the model is less than the number of observed variables. The Gali and Monacelli model has only two stochastic shocks, which is fewer than the data used here.

The endogenous part of the model consists of four equations: a forward looking IS equation, a Phillips curve, an exchange rate policy equation and a monetary policy rule. The open economy IS curve is derived from the consumption Euler equation:

$$\begin{aligned} y_t = {} & E_t y_{t+1} - \left[\tau + \alpha(2-\alpha)(1-\tau)\right](R_t - E_t \pi_{t+1}) - \rho_z z_t \\ & - \alpha\left[\tau + \alpha(2-\alpha)(1-\tau)\right] E_t \Delta q_{t+1} + \alpha(2-\alpha)\frac{1-\tau}{\tau} E_t \Delta y^*_{t+1}, \end{aligned} \tag{2.1}$$

where τ is the inter-temporal substitution elasticity, $0 < \alpha < 1$ the import share, y_t aggregate output, π_t inflation rate, and q_t terms of trade. Furthermore, z_t is the growth rate of a non-stationary world technology process, A_t and y_t^* is exogenous world output. Notice that all real variables are expressed in terms of percentage deviation from A_t.

Optimal price setting by monopolistic competitors who produce intermediate goods leads to the forward-looking Phillips curve:

$$\pi_t = \beta E_t \pi_{t+1} + \alpha\beta E_t \Delta q_{t+1} - \alpha \Delta q_t + \frac{\kappa}{\tau + \alpha(2-\alpha)(1-\tau)}\left(y_t - \bar{y}_t\right), \tag{2.2}$$

where $\bar{y}_t = -\alpha(2-\alpha)(1-\tau)/\tau y_t^*$ is potential output in the absence of nominal rigidities, β the discount factor and $\kappa > 0$ is the degree of price stickiness. Assuming that purchasing power parity holds, we can set up the nominal exchange rate e_t from the definition of the CPI:

$$\pi_t = \Delta e_t + (1-\alpha)\Delta q_t + \pi_t^* \tag{2.3}$$

where π_t^* is a world inflation shock. Monetary policy is described by a Taylor-type interest rate rule, where the interest rate responds to inflation, output and nominal exchange rate depreciation:

$$R_t = \rho_R R_{t-1} + (1-\rho_R)\left[\psi_1 \pi_t + \psi_2\left(y_t - \bar{y}_t\right) + \psi_3 \Delta e_t\right] + \varepsilon_t^R, \tag{2.4}$$

where ρ_R is a smoothing term and ε_t^R an exogenous policy shock. We assume that the policy coefficients ψ_1, ψ_2, ψ_3 are ≥ 0. The law of motion for the terms of trade is exogenously given:

$$\Delta q_t = \rho_q \Delta q_{t-1} + \varepsilon_t^q. \tag{2.5}$$

Equations (2.1)-(2.5) form a linear rational expectation model. Moreover, we assume that world output and inflation, y_t^* and π_t^* follow exogenous autoregressive processes:

$$\pi_t^* = \rho_{\pi *}\pi_{t-1}^* + \varepsilon_t^{\pi *}, \quad y_t^* = \rho_{y*}y_{t-1}^* + \varepsilon_t^{y*}. \tag{2.6}$$

4.3 Econometric model: Bayesian estimation

There are numerous econometric estimation procedures such as calibration, generalized method of moments (GMM), minimum distance approach, full information maximum likelihood estimation and Bayesian approach. This paper uses the Bayesian estimation approach for DSGE models, which has recently gained some success in macroeconomic modelling[8].

There are several practical reasons to use the Bayesian approach. First, compared with maximum likelihood and GMM estimators, the Bayesian approach does not suffer from small sample issues. For instance, Christiano and den Haan (1996) have discussed that in a small sample, GMM estimates of a DSGE model often have a distribution which is not implied by asymptotic theory. Second, according to Fernandez-Villaverde and Rubio-Ramirez (2004), Bayesian estimation and model comparison are consistent in the case of mis-specified models. Third, there are often strong restrictions on the structural parameters of a DSGE model that are not easy to impose with GMM and maximum likelihood. Fourth, the Bayesian approach offers a set of answers that are relevant for policymakers who want to know, conditional on the observed data, the probability of pursuing the right policy (Fernandez-Villaverde 2010).

Bayesian estimation of a DSGE model consists of the following four steps. First, we solve the DSGE model by linearizing around a steady state using a first-order Taylor series expansion and rewrite the solution to a DSGE model as a system of Kalman filter equations:

[8] For example, see Schorfheide 2000, Smets and Wouters 2003 and An and Schorfheide 2005.

$$y_t^* = M\bar{y}(\theta) + M\hat{y}_t + N(\theta)x_t + \eta_t \tag{3.1}$$

$$\hat{y}_t = g_y(\theta)\hat{y}_{t-1} + g_u(\theta)u_t \tag{3.2}$$

$$E(\eta_t\eta_t') = V(\theta) \tag{3.3}$$

$$E(u_tu_t') = Q(\theta) \tag{3.4}$$

where $\hat{y}_t$ is the vector of variables in deviations from steady state, $\bar{y}$ is the vector of steady state values and θ, the vector of structural parameters to be estimated. Equation (3.1) is a measurement equation which expresses the relationship between observed variables (y_t^*) and model variables ($\hat{y}_t$) with measurement error (η_t). In order to cope with stochastic singularities, the measurement error terms are added to the measurement equation. As already noted, stochastic singularity problems arise if the number of shock terms in the DSGE model is less than the number of observed variables. For instance, the one shock assumption makes the real business cycle model stochastically singular, and the model predicts that certain combinations of endogenous variables will be deterministic. But such an exact relationship usually does not hold in practice. Moreover, model variables may have a trend, which is captured by $N(\theta)x_t$. Equation (3.2) is a transition equation that comes from the DSGE model solution. $V(\theta)$ and $Q(\theta)$ are the variances of measurement errors and structural shocks.

Second, we estimate the likelihood of the DSGE model using a Kalman filter recursion. For $t = 0,\dots,T$ and with initial values y_0 and P_0 given, the Kalman filter recursion is as follows:

$$v_t = y_t^* - \bar{y}^* - M\hat{y}_t - Nx_t \tag{3.5}$$

$$F_t = MP_tM' + V \tag{3.6}$$

$$K_t = g_yP_tg_y'F_t^{-1} \tag{3.7}$$

$$\hat{y}_{t+1} = g_y \hat{y}_t + K_t v_t \tag{3.8}$$

$$P_{t+1} = g_y P_t (g_y - K_t M)' + g_u Q g_u' \tag{3.9}$$

A log-likelihood function is derived from this filter recursion:

$$\ln L(\tilde{\theta} \mid Y_t^*) = -\frac{Tk}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} |F_t| - \frac{1}{2} v_t' F_t^{-1} v_t \tag{3.10}$$

where the vector $\tilde{\theta}$ contains the deep parameters , θ, variances , $V(\theta)$ and $Q(\theta)$, and

where Y_t^* expresses the set of observable endogenous variables y_t^* found in the

measurement equation.[9] The log posterior kernel can be written as

$$\ln K(\tilde{\theta} \mid Y_t^*) = \ln L(\tilde{\theta} \mid Y_t^* + \ln p(\tilde{\theta}) \tag{3.11}$$

where $\ln p(\tilde{\theta})$ are the prior distributions.

Third, we maximize the log posterior kernel with respect to the structural parameters θ and find the posterior mode of the log posterior kernel.

Finally, we estimate the posterior distribution of the parameters using a Markov chain Monte Carlo algorithm. The log posterior kernel is a nonlinear and complicated function of the parameter, θ. Thus it is impossible to calculate analytically the conditional expected value of the parameter θ. Instead we use a numerical integration procedure, the Metropolis-Hasting algorithm, to find it.[10]

The random walk Metropolis-Hastings algorithm consists of the following steps:

1. Choose a starting parameter of value $\tilde{\theta}^\circ$, where this is typically the posterior mode which maximizes the log posterior kernel (3.11) and a loop over step 2-3-4
2. Draw a proposal $\tilde{\theta}^*$ from a normal distribution (candidate or jumping distribution)

$$J(\tilde{\theta}^* \mid \tilde{\theta}^{t-1}) = N(\tilde{\theta}^{t-1}, c\Sigma_m) \tag{3.12}$$

[9] A good overview of the Kalman filter is provided by Hamilton (1994).
[10] For a good introduction, see Chib and Greenberg (1995).

where Σ_m is the inverse of the Hessian computed at the posterior mode and c is a scale factor.

3. Compute the acceptance ratio

$$r = \frac{p(\tilde{\theta}^* \mid Y_T^*)}{p(\tilde{\theta}^{t-1} \mid Y_T^*)} = \frac{K(\tilde{\theta}^* \mid Y_T^*)}{K(\tilde{\theta}^{t-1} \mid Y_T^*)} \tag{3.13}$$

4. Finally accept or reject the proposal $\tilde{\theta}^*$ according to the following rule

$$\tilde{\theta}^t = \begin{cases} \tilde{\theta}^* & \textit{with probability} \min(r,1) \\ \tilde{\theta}^{t-1} & \textit{otherwise} \end{cases} \tag{3.14}$$

4.4 Data and prior description

Monetary policy reaction functions were estimated using quarterly data from 1997:1 to 2008:4. We estimated the reaction function using four variables: output growth, inflation (change in CPI), exchange rate changes and the nominal interest rate. The consumer price index and real GDP (in constant prices of 2005) are taken from the bulletin of the National Statistical Office of Mongolia. Data for the exchange rate (togrog against US dollar) and the central bank bill rate are from the bulletin of the Bank of Mongolia (BOM). All variables except the interest rate are in logarithms and seasonally adjusted using the Census X12 approach. Output growth, inflation, and exchange rate changes are computed as log differences of the respective raw series. In order to obtain percentage changes we multiply output growth and exchange rates by 100 and inflation by 400. Furthermore, all variables are de-meaned before estimation.

Table 4.1 shows the prior distribution of the model parameters. Following a Taylor rule, the priors for ψ_1 and ψ_2 are set at values of 1.5 and 0.25 respectively. The prior mean (ψ_3) is chosen at 0.25 for the exchange rate parameter. The model is parameterized in terms of the steady state real interest rate instead of the discount factor ($\beta = \exp[-r/400]$). Its mean is set at 2.5% with a large standard deviation.

Table 4.1 Prior distributions

Name	Domain	Density	Density parameters	
			P(1)	P(2)
ψ_1	$\mathbb{R}^+$	Gamma	1.50	0.50
ψ_2	$\mathbb{R}^+$	Gamma	0.25	0.13
ψ_3	$\mathbb{R}^+$	Gamma	0.25	0.13
ρ_R	$[0,1)$	Beta	0.50	0.20
α	$[0,1)$	Beta	0.60	0.10
r	$\mathbb{R}^+$	Gamma	2.50	1.00
κ	$\mathbb{R}^+$	Gamma	0.50	0.25
τ	$[0,1)$	Beta	0.50	0.20
ρ_q	$[0,1)$	Beta	0.80	0.20
ρ_z	$[0,1)$	Beta	0.40	0.05
ρ_{y*}	$[0,1)$	Beta	0.70	0.10
$\rho_{\pi*}$	$[0,1)$	Beta	0.70	0.17
σ_R	$\mathbb{R}^+$	InvGamma	1.25	0.65
σ_q	$\mathbb{R}^+$	InvGamma	2.50	1.31
σ_z	$\mathbb{R}^+$	InvGamma	1.25	0.65
σ_{y*}	$\mathbb{R}^+$	invGamma	1.20	0.65
$\sigma_{\pi*}$	$\mathbb{R}^+$	InvGamma	1.88	0.98

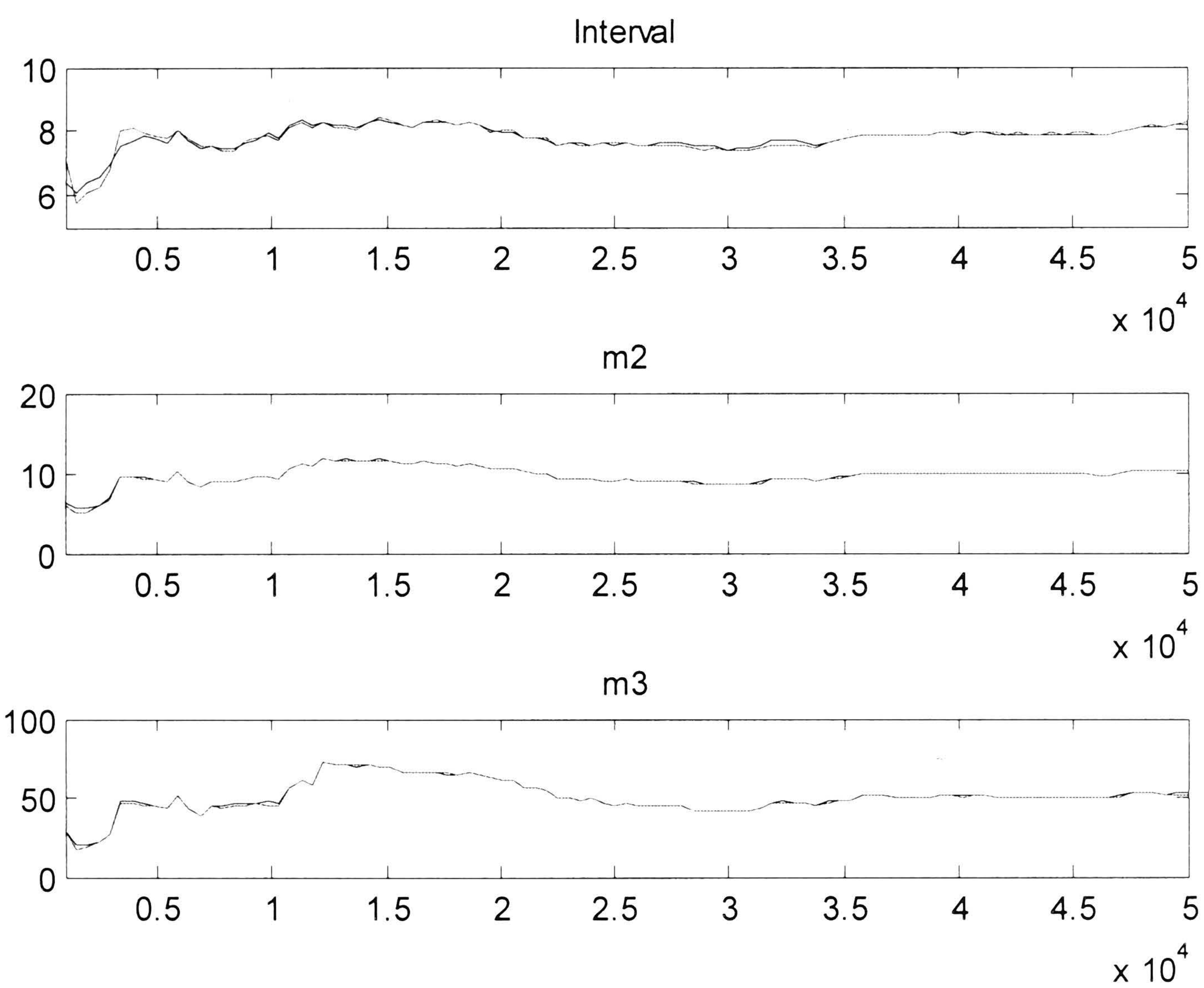
Notes: P(1) and P(2) are the means and standard deviations for the beta and gamma distributions; the upper and lower bound of the uniform distribution; s and v for the inverse gamma distribution ($p(\sigma \mid v, s) \propto \sigma^{-v-1} e^{-vs^2/2\sigma^2}$).
Source: Author's calculations

The mean of the slope coefficient (κ) in the Phillips curve is chosen to be 0.5 with a standard deviation of 0.25. Since the import/GDP ratio is roughly 60 percent, the prior for import share coefficient is tightly centered at 0.6. Also we restrict $0 < \tau < 1$ with a prior mean of 0.5. To specify the priors for the exogenous shock process, we use an autoregressive model of order 1.

4.5 Estimation results

The Bayesian estimation of the model is shown in Table 4.2. First, on the basis of the independent prior distribution, the posterior mode is found using Sims' minimization algorithm which is robust against the cliff. Then starting from these modes, we estimate the parameters by drawing from a random walk Metropolis-Hastings algorithm with 50000 replications, two parallel chains and the scale factor $c = 0.5$. The acceptance rate was approximately 28 percent for two chains. Convergence was tested using the Brooks and Gelmen (1998) statistics (Figure 4.1).

Figure 4.1 Multivariate diagnostic

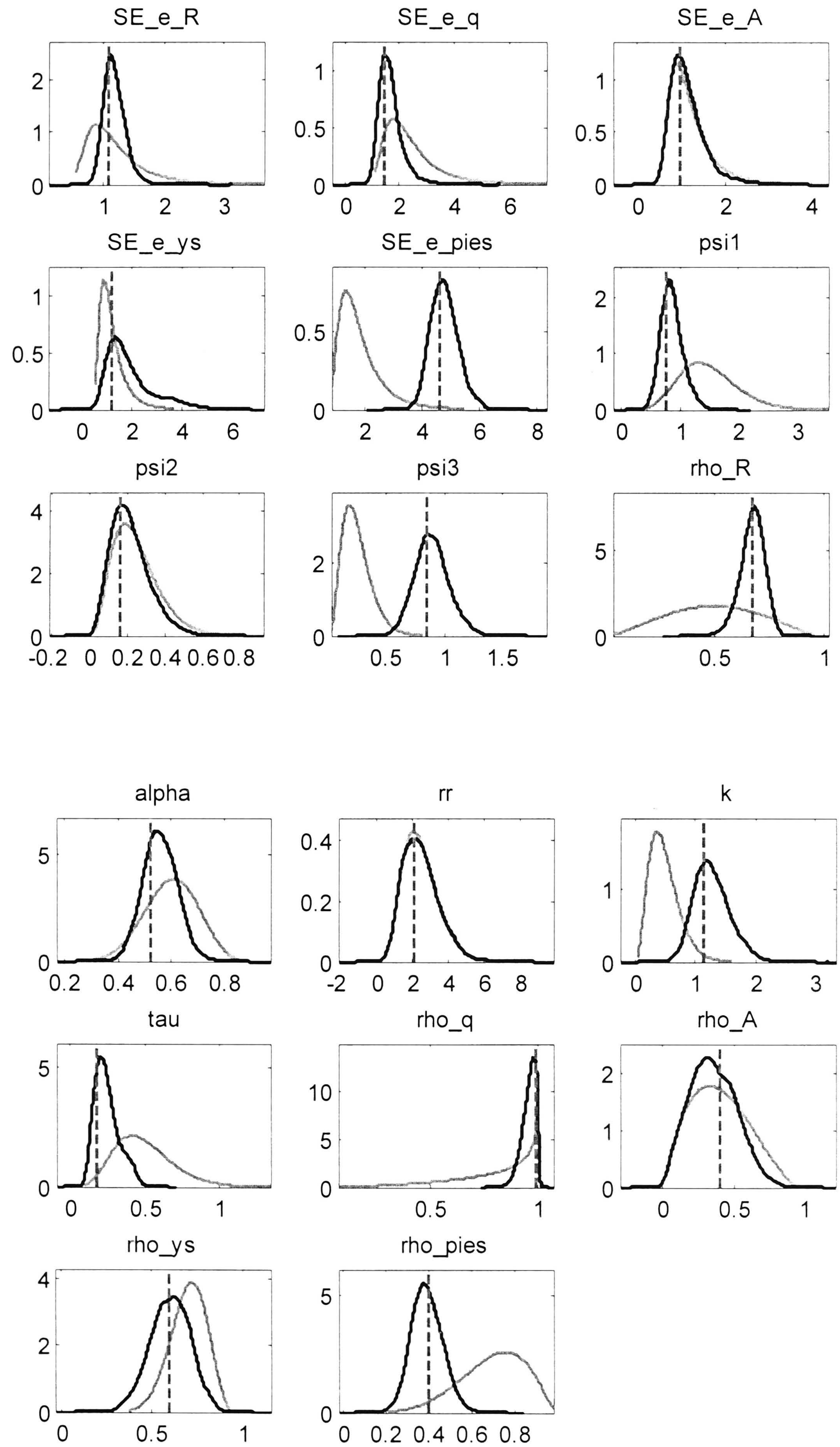


Source: Author's calculations

The three measures of interval, a measure constructed from an 80% confidence interval around the parameter mean, m2, a measure of the variance and m3, a measure based on third moments all are relatively constant and convergent. The Markov chain Monte Carlo univariate diagnostics is provided in Figure 4.5a of the appendix.

In Figure 4.2 we present the posterior distribution (black line) of the parameters with the prior distribution (grey line) and posterior mode (dashed line). It is observed that the posterior distribution is close to a normal distribution around the posterior mode.

Figure 4.2 Posterior distribution



Source: Author's calculations

The posterior means for policy coefficients of inflation and exchange rate differ markedly from their priors. We find that the Bank of Mongolia pursues a reasonably anti-inflationary policy ($\psi_1 = 0.86$), demonstrates concern for output ($\psi_2 = 0.21$) and responds to the exchange rate very actively ($\psi_3 = 0.89$) for a simple Taylor rule. Furthermore, the degree of interest rate smoothing is relatively high with an estimate of $\rho_R = 0.67$.

Table 4.2 Parameter estimation results

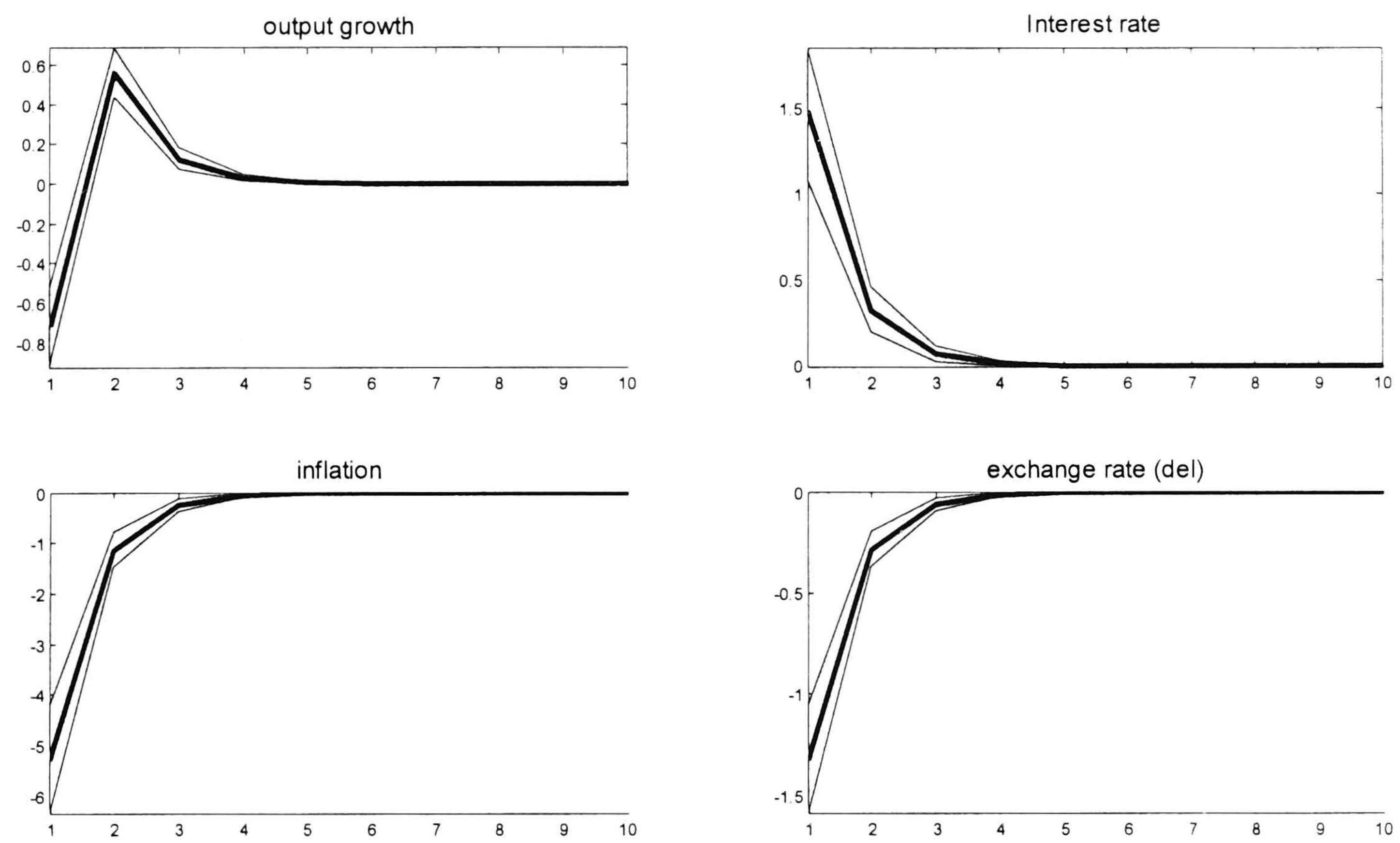
Parameters	Prior mean	Posterior mode	Posterior mean	Confidence interval (90%)	
ψ_1	1.50	0.78	0.86	0.56	1.15
ψ_2	0.25	0.16	0.21	0.05	0.37
ψ_3	0.25	0.86	0.89	0.65	1.13
ρ_R	0.50	0.67	0.67	0.59	0.77
α	0.60	0.52	0.56	0.46	0.66
r	2.50	2.09	2.48	0.94	4.02
κ	0.50	1.15	1.28	0.81	1.77
τ	0.50	0.17	0.24	0.12	0.38
ρ_q	0.80	0.99	0.96	0.91	1.00
ρ_z	0.40	0.40	0.35	0.08	0.61
ρ_{y*}	0.70	0.59	0.60	0.42	0.77
$\rho_{\pi*}$	0.70	0.40	0.39	0.28	0.51
σ_R	1.25	1.08	1.18	0.90	1.45
σ_q	2.51	1.46	1.66	1.05	2.28
σ_z	1.25	0.98	1.14	0.57	1.68
σ_{y*}	1.25	1.20	1.90	0.72	3.45
$\sigma_{\pi*}$	1.88	4.64	4.77	3.98	5.50

Source: Author's calculations

Other structural coefficients, such as import share α, slope of Phillips curve κ, and intertemporal substitution elasticity τ, are within reasonable ranges. The estimates of the law of motion for the terms of trade show a very high degree of persistence compared with other stochastic processes.

In order to explore the dynamic effect of the reaction function shock we compute impulse response functions. These are illustrated in Figure 4.3.

Figure 4.3 Impulse response function for monetary shock



Source: Author's calculations

Contractionary monetary policy leads to an appreciation of the exchange rate and a lower inflation and output. The impulse response functions for other shocks and variance decompositions are provided in Figure 4.6a and Table 4.7a of the appendix.

We can test the output and exchange rate coefficients of the policy reaction function by calculating the posterior odds ratio. The test results are presented in Table 4.3. According to the posterior odds ratio test, hypothesis $\psi_2 = 0$ against the alternative $\psi_2 > 0$ is not rejected, while hypothesis $\psi_3 = 0$ against the alternative $\psi_3 > 0$ is rejected.

Table 4.3 Posterior odds

Log marginal data densities (modified harmonic mean)		Odds
$\psi_2 = 0$	$\psi_2 > 0$	
-608.4363	-608.8664	1.5374
$\psi_3 = 0$	$\psi_3 > 0$	
-633.6108	-608.8664	0.0000

Source: Author's calculations

We also test the robustness of the specification of the monetary policy rule using the marginal data density. The model estimation under a simple output rule rather than the output gap rule is not shown in the paper for the sake of brevity. According to the posterior odds ratio, a simple output rule is rejected. The model estimation under an expected inflation rule is shown in Table 4.4. We found that a forward-looking inflation rule improves the model fit, as measured by the marginal data densities, compared with the benchmark specification (log data density is -604.84). More importantly, the policy response to the expected inflation increases significantly and the Taylor principle holds.

Table 4.4 Expected inflation rule

Parameters	Prior mean	Posterior mean	Confidence interval (90%)	
ψ_1	1.50	1.59	1.07	2.08
ψ_2	0.25	0.23	0.06	0.40
ψ_3	0.25	0.73	0.49	0.95
ρ_R	0.50	0.57	0.45	0.68
α	0.60	0.59	0.49	0.69
r	2.50	2.47	0.96	3.90
κ	0.50	1.32	0.85	1.78
τ	0.50	0.24	0.13	0.34
ρ_q	0.80	0.96	0.93	1.00
ρ_z	0.40	0.32	0.06	0.54
ρ_{y*}	0.70	0.61	0.43	0.79
$\rho_{\pi*}$	0.70	0.35	0.24	0.45
σ_R	1.25	1.04	0.81	1.24
σ_q	2.51	1.39	0.97	1.77
σ_z	1.25	1.19	0.59	1.83
σ_{y*}	1.25	1.70	0.76	2.63
$\sigma_{\pi*}$	1.88	4.69	3.88	5.41

Source: Author's calculations

4.6 Forecasting

Following DeJong et al. (1993), Ingram and Whiteman (1994) and Del Negro and Schorfheide (2004), we can use a DSGE model as a prior for a vector autoregression. Del Negro and Schorfheide priors differ from the one used by DeJong et al. (1993) who used a simulation procedure to approximate the marginal prior for the VAR coefficients by a conjugate Inverted-Wishart Normal (IW-N) prior.

In this paper we follow Del Negro and Schorfheide's approach to forecast inflation and output. The idea is to make the DSGE model prior by generating artificial observations from the DSGE model and adding these dummy observations to actual data of the VAR model. Consider the order p VAR representation for $n \times 1$ vector y_t of observed variables:

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + u_t \tag{6.1}$$

where innovations $u_t \sim N(0, \Sigma_u)$.Let Y be the $T \times n$ matrix with rows y_t'. Let $k = 1 + np$, X the $T \times k$ matrix with rows $x_t' = \left[1, y_{t-1}', \ldots, y_{t-p}'\right]$, U the $T \times n$ matrix with rows u_t', and $\Phi = \left[\Phi_0, \Phi_1, \ldots, \Phi\right]'$. The VAR can be rewritten as $Y = X\Phi + U$ with likelihood function:

$$p(Y \mid \Phi, \Sigma_u) \propto |\Sigma_u|^{-T/2} \exp\left\{-\frac{1}{2} tr\left[\Sigma_u^{-1}\left(Y'Y - \Phi'X'Y - Y'X\Phi + \Phi'X'X\Phi\right)\right]\right\} \tag{6.2}$$

conditional on observations $y_{1-p}, \ldots, y_0$. Suppose the actual observations are joined with $T^* = \lambda T$ dummy observations $\left(Y^*, X^*\right)$ generated from the DSGE model. The likelihood function for the combined sample is obtained by premultiplying (6.2) with

$$\begin{aligned} & p(Y^*(\theta) \mid \Phi, \Sigma_u) \\ & \propto |\Sigma_u|^{-T/2} \exp\left\{-\frac{1}{2} tr\left[\Sigma_u^{-1}\left(Y^{*\prime}Y^* - \Phi'X^{*\prime}Y^* - Y^{*\prime}X^*\Phi + \Phi'X^{*\prime}X^*\Phi\right)\right]\right\} \end{aligned} \tag{6.3}$$

In order to remove stochastic variation, Del Negro and Schorfheide (2004) use the DSGE theoretical autocovariance matrices, $\lambda T\Gamma^*_{yy}(\theta)$, $\lambda T\Gamma^*_{yx}(\theta)$ and $\lambda T\Gamma^*_{xx}(\theta)$ instead of the sample moments $Y^{*\prime}Y^*$, $Y^{*\prime}X^*$, and $X^{*\prime}X^*$.

Conditional on θ the prior distribution $p(\Phi, \Sigma_u \mid \theta)$ of the VAR parameters is of the IW-N form:

$$\Sigma_u \mid \theta \sim IW(\lambda T\Sigma^*_u, \lambda T - k, n) \tag{6.4}$$

$$\Phi \mid \Sigma_u, \theta \sim N\left(\Phi^*(\theta), \Sigma_u \otimes \left(\lambda T\Gamma^*_{xx}(\theta)\right)^{-1}\right) \tag{6.5}$$

where $\Phi^*(\theta) = \Gamma^{*-1}_{xx}(\theta)\Gamma^*_{xy}(\theta)$, $\Sigma^*_u(\theta) = \Gamma^*_{yy}(\theta) - \Gamma^*_{yx}(\theta)\Gamma^{*-1}_{xx}(\theta)\Gamma^*_{xy}(\theta)$. The full specification of the prior is completed with the prior distribution of parameters $p(\theta)$:

$$p(\Phi, \Sigma, \theta) = p(\Phi, \Sigma \mid \theta)\, p(\theta). \tag{6.6}$$

The posterior distribution can be factorized into the posterior density of the VAR parameters $p(\Phi, \Sigma_u \mid Y, \theta)$ and the marginal posterior density of the DSGE model parameters $p(\theta \mid Y)$:

$$p(\Phi, \Sigma_u, \theta \mid Y) = p(\Phi, \Sigma_u \mid Y, \theta)\, p(\theta \mid Y) \tag{6.7}$$

Following Zellner (1971) it can be shown that conditional on θ, the posterior distribution of Φ and Σ is also of the IW-N form:

$$\Sigma_u \mid Y, \theta \sim IW\left((\lambda+1)T\tilde{\Sigma}_u(\theta), (1+\lambda)T - k, n\right) \tag{6.8}$$

$$\Phi \mid Y, \Sigma_u, \theta \sim N\left(\tilde{\Phi}(\theta), \Sigma_u \otimes \left(\lambda T\Gamma^*_{xx}(\theta) + X'X\right)^{-1}\right) \tag{6.9}$$

where $\tilde{\Phi}(\theta) = \left(\lambda T\Gamma^*_{xx}(\theta) + X'X\right)^{-1}\left(\lambda T\Gamma^*_{xy} + X'Y\right)$,

$\tilde{\Sigma}_u(\theta) = \frac{1}{(\lambda+1)T}\Big[\left(\lambda T\Gamma_{yy} + Y'Y\right) -$

$-\left(\lambda T\Gamma^{*}_{yx}(\theta)+Y'X\right)\left(\lambda T\Gamma^{*}_{xx}(\theta)+X'X\right)^{-1}\left(\lambda T\Gamma^{*}_{xy}(\theta)+X'Y\right)\Big]$. The hyper-parameter λ is chosen to maximise the marginal data density:

$$p_{\lambda}(Y)=\int p_{\lambda}(Y\,|\,\theta)\,p(\theta)\,d\theta. \tag{6.10}$$

We use Geweke's (1999) modified harmonic mean estimator to obtain numerical approximation of the marginal data densities over a grid that contains values of $\lambda=\{0.5,1.0,1.5,2.5,10,\text{inf}\}$ and find that the optimal λ is 1.5. The estimation result is shown in Table 4.5.

Table 4.5 Optimal λ

	λ					
	0.5	1.0	1.5	2.5	10	inf
Log of the marginal data density	-600.02	-584.16	-581.77	-582.44	-590.61	-596.41

Source: Author's calculations

The value of the optimal λ implies weights of 40 percent on the VAR and 60 percent on the DSGE model. The structural parameter estimates of the posterior simulation are shown in Table 4.6 and forecasts of output, inflation, interest rate and the exchange rate are presented in Figure 4.4.

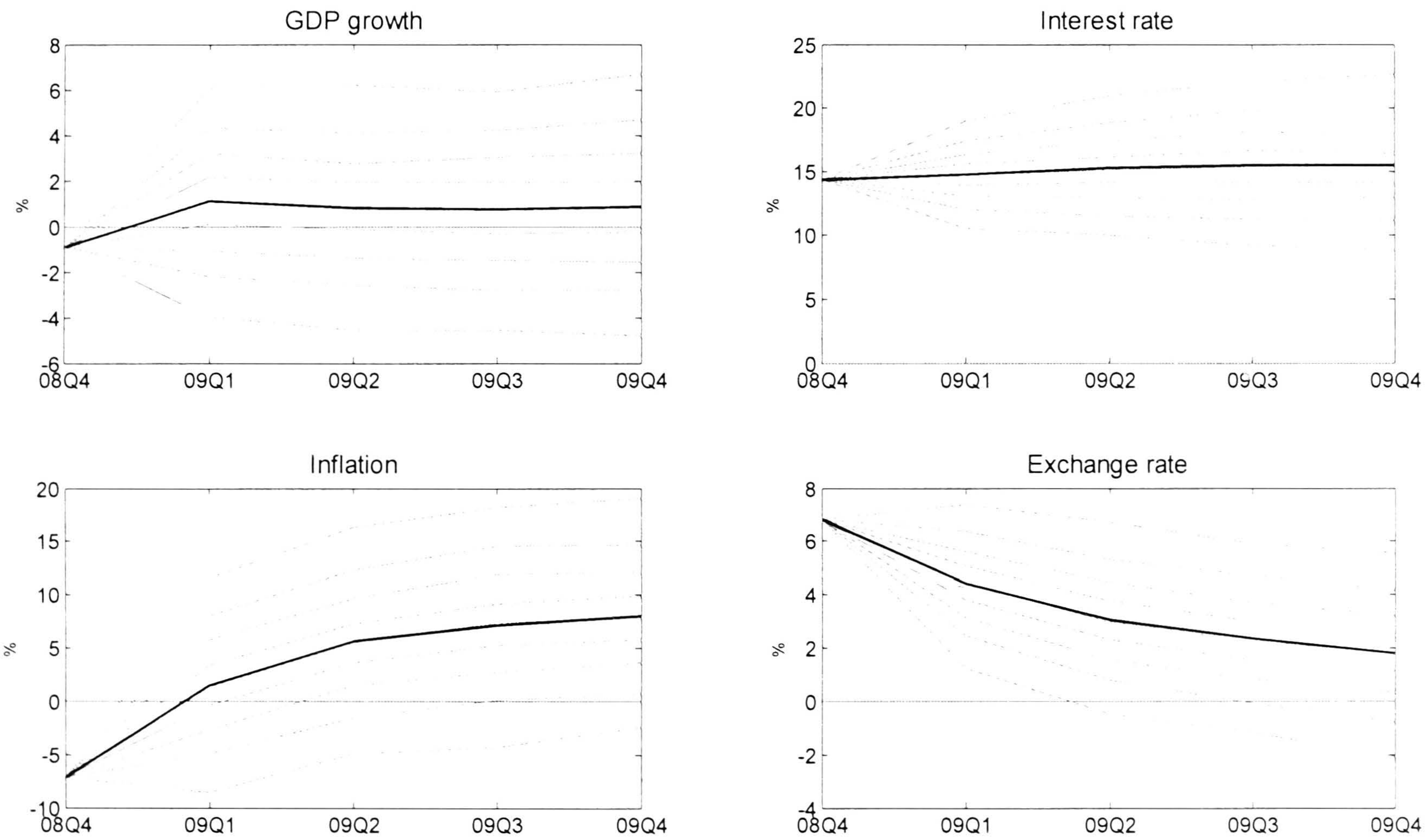
Table 4.6 DSGE-VAR estimation

Parameters	Prior mean	Posterior mean	Confidence interval (90%)	
ψ_1	1.50	1.05	0.64	1.41
ψ_3	0.25	0.69	0.43	0.96
ρ_R	0.50	0.64	0.52	0.77
α	0.60	0.62	0.49	0.74
r	2.50	2.44	0.86	3.85
κ	0.50	0.78	0.52	1.03
τ	0.50	0.30	0.17	0.42
ρ_q	0.80	0.82	0.68	0.96
ρ_z	0.40	0.29	0.05	0.53
ρ_{y*}	0.70	0.65	0.48	0.83

$\rho_{\pi*}$	0.70	0.54	0.41	0.69
σ_R	1.25	0.92	0.69	1.15
σ_q	2.51	1.55	0.98	2.09
σ_z	1.25	1.02	0.55	1.48
σ_{y*}	1.25	1.50	0.65	2.32
$\sigma_{\pi*}$	1.88	2.93	2.27	3.60

Source: Author's calculation

Figure 4.4 Out-of-sample forecasting



Source: Author's calculations

From Figure 4.4, we can see that on average the inflation is expected to be 8 percent while output increases by 4 percent at the end of 2009.

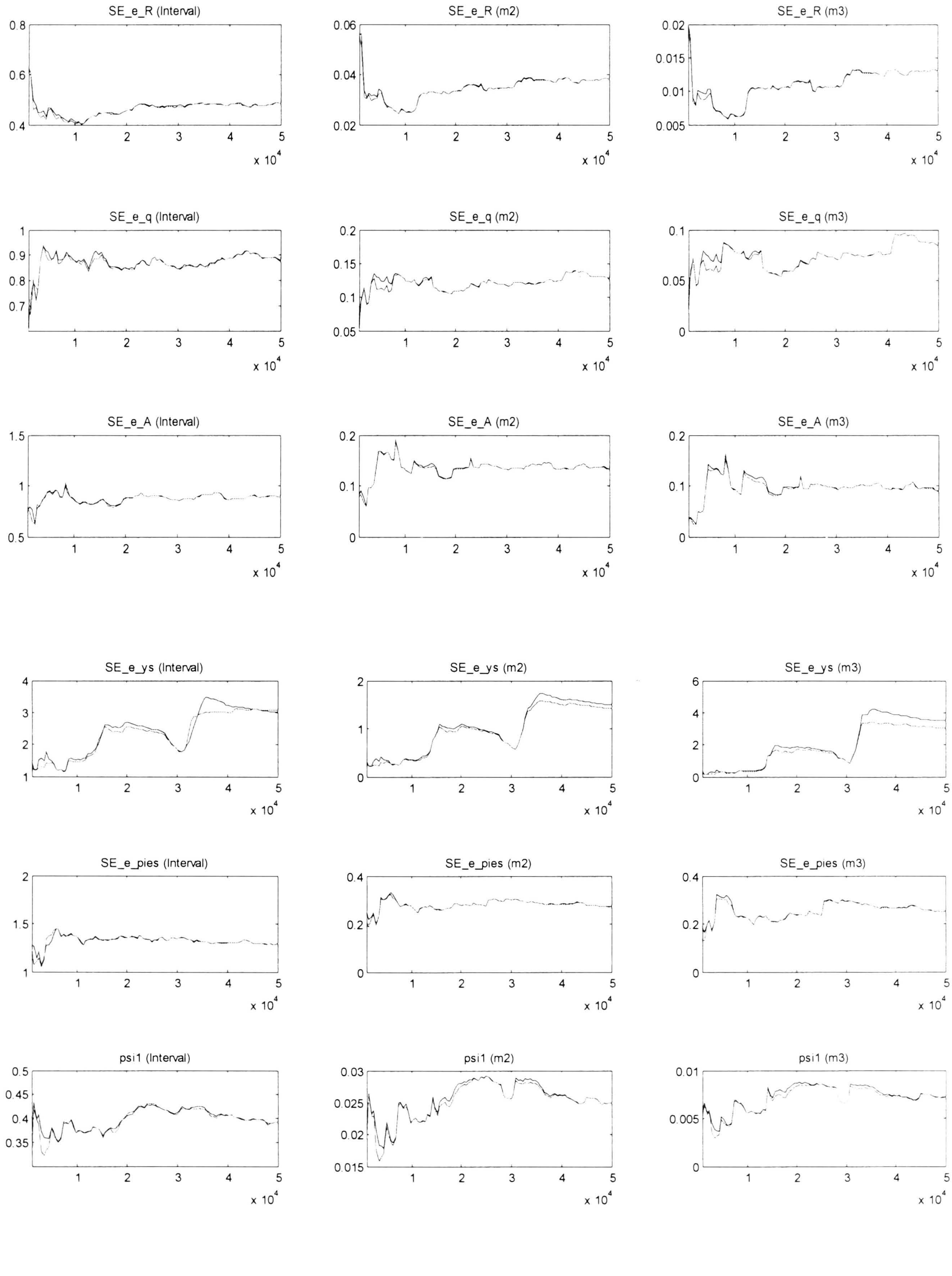
4.7 Conclusions

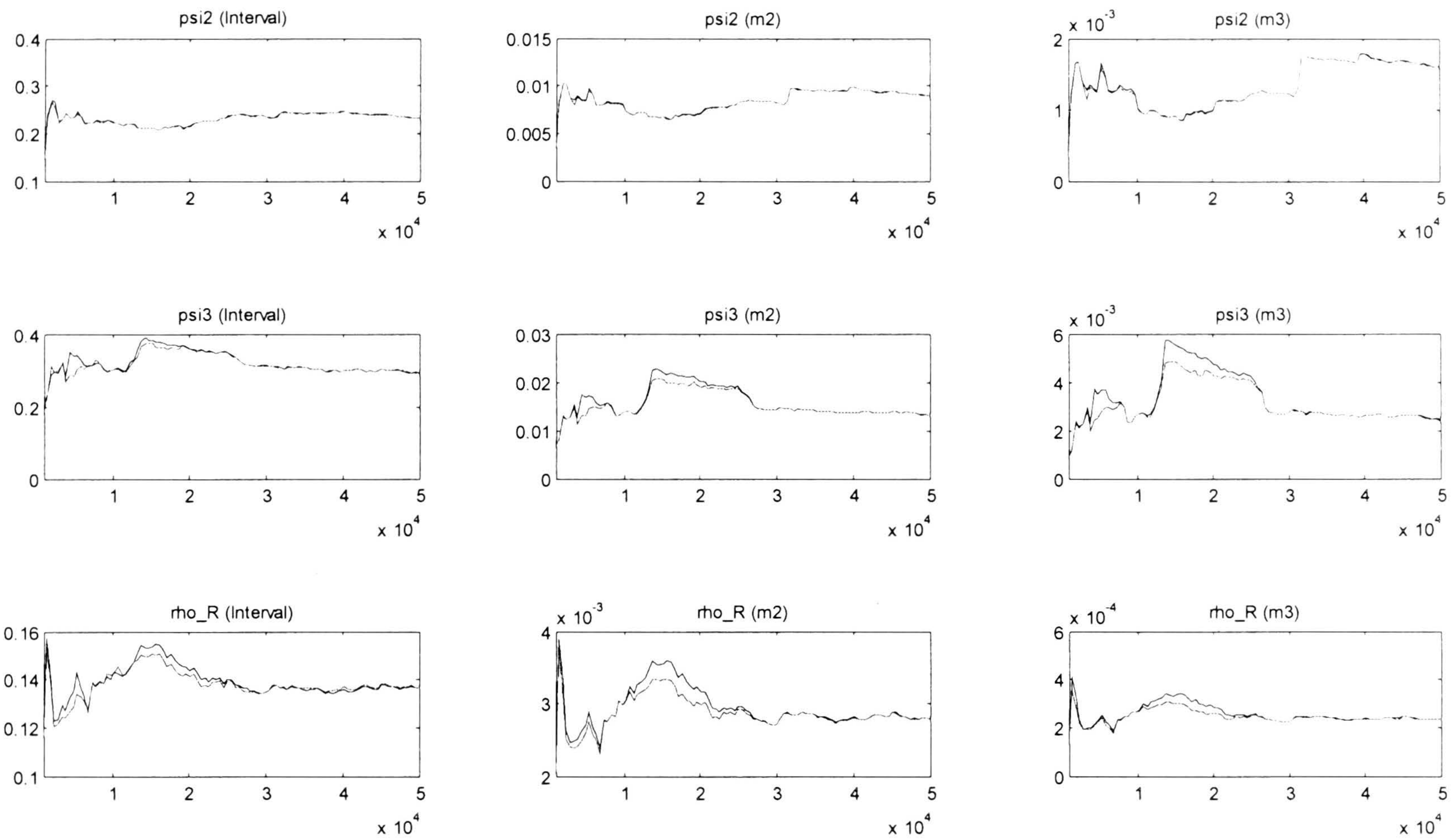
We estimate the monetary policy reaction function using the New Keynesian small open economy DSGE model developed by Gali and Monacelli (2005) and modified for estimation purposes by Lubik and Schorfheide (2007) on Mongolian data for 1997-2008. We also estimate the DSGE-VAR model for forecasting purposes.

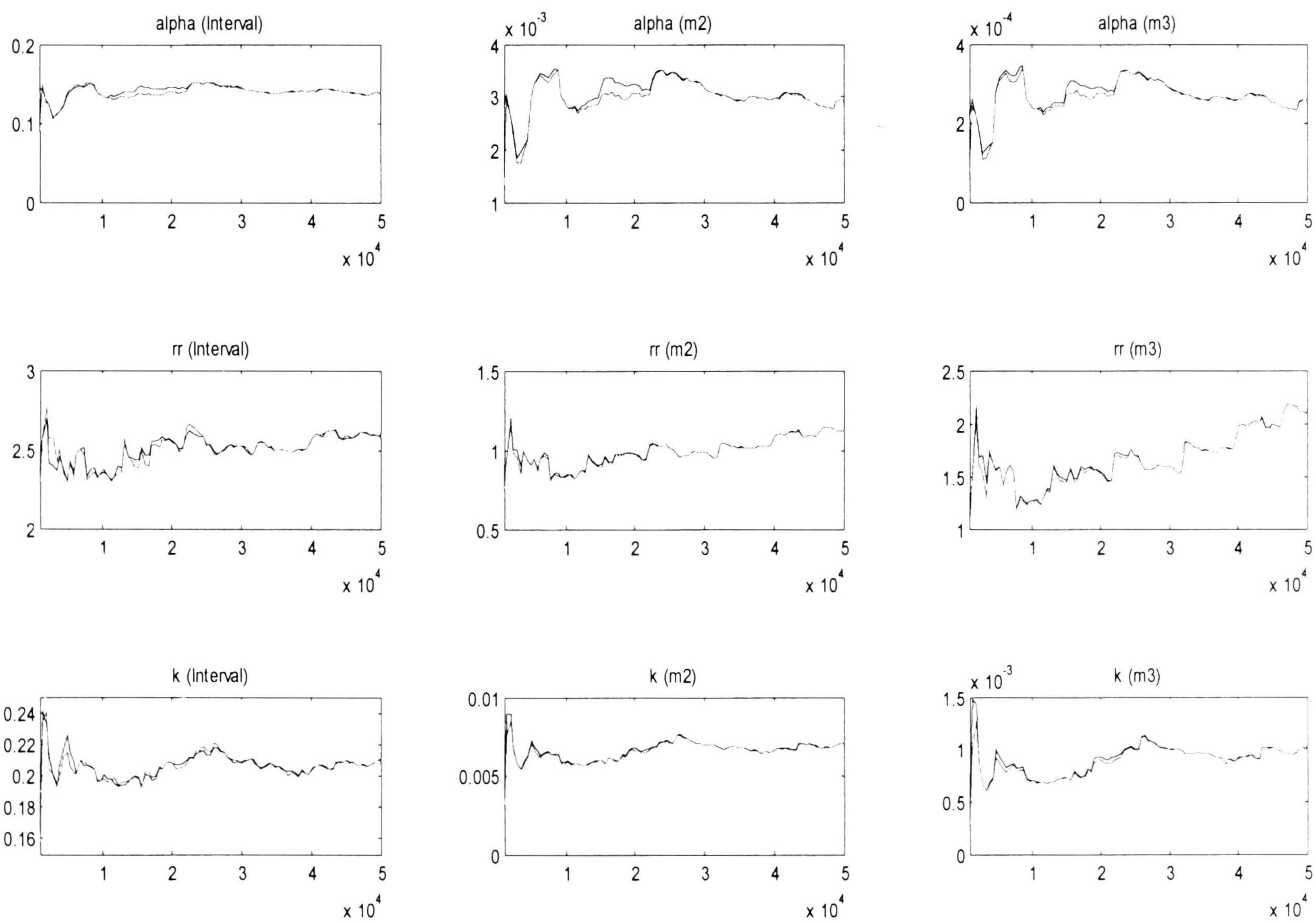
Our main findings are summarized as follows. First, the monetary policy reaction function is forward-looking in terms of the inflation rate. The expected inflation rule fits the reaction function better than a simple Taylor type rule. Second, the central bank of Mongolia has implemented a strong anti-inflationary and exchange rate stabilization policy. Third, there is evidence that the Bank of Mongolia does not respond significantly to output according to the Bayesian posterior odds. Finally, the DSGE-VAR analysis demonstrates that our estimated small open economy model can be used for forecasting output and inflation.

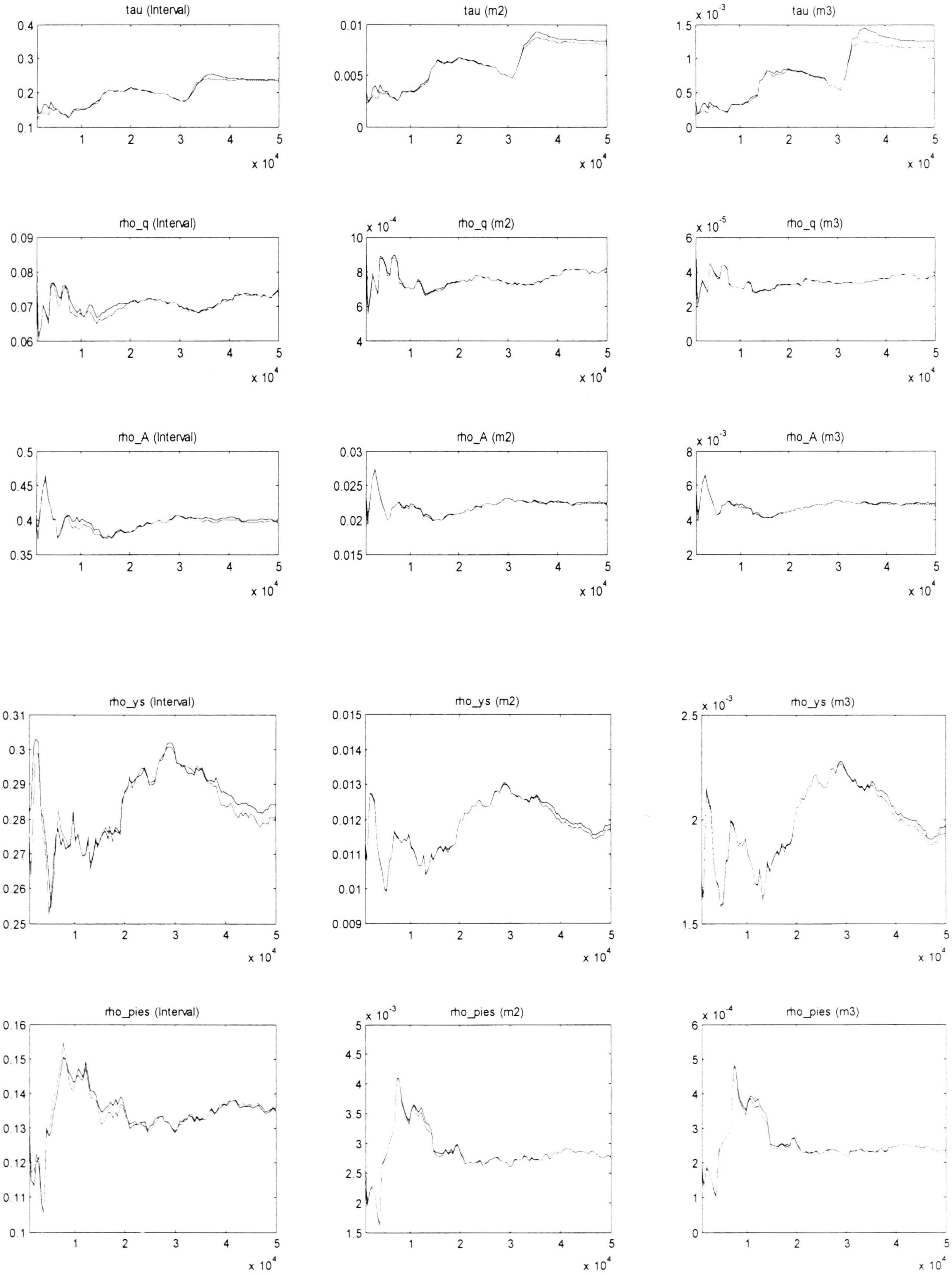
4.8 Appendix

Figure 4.5a MCMC univariate diagnostic



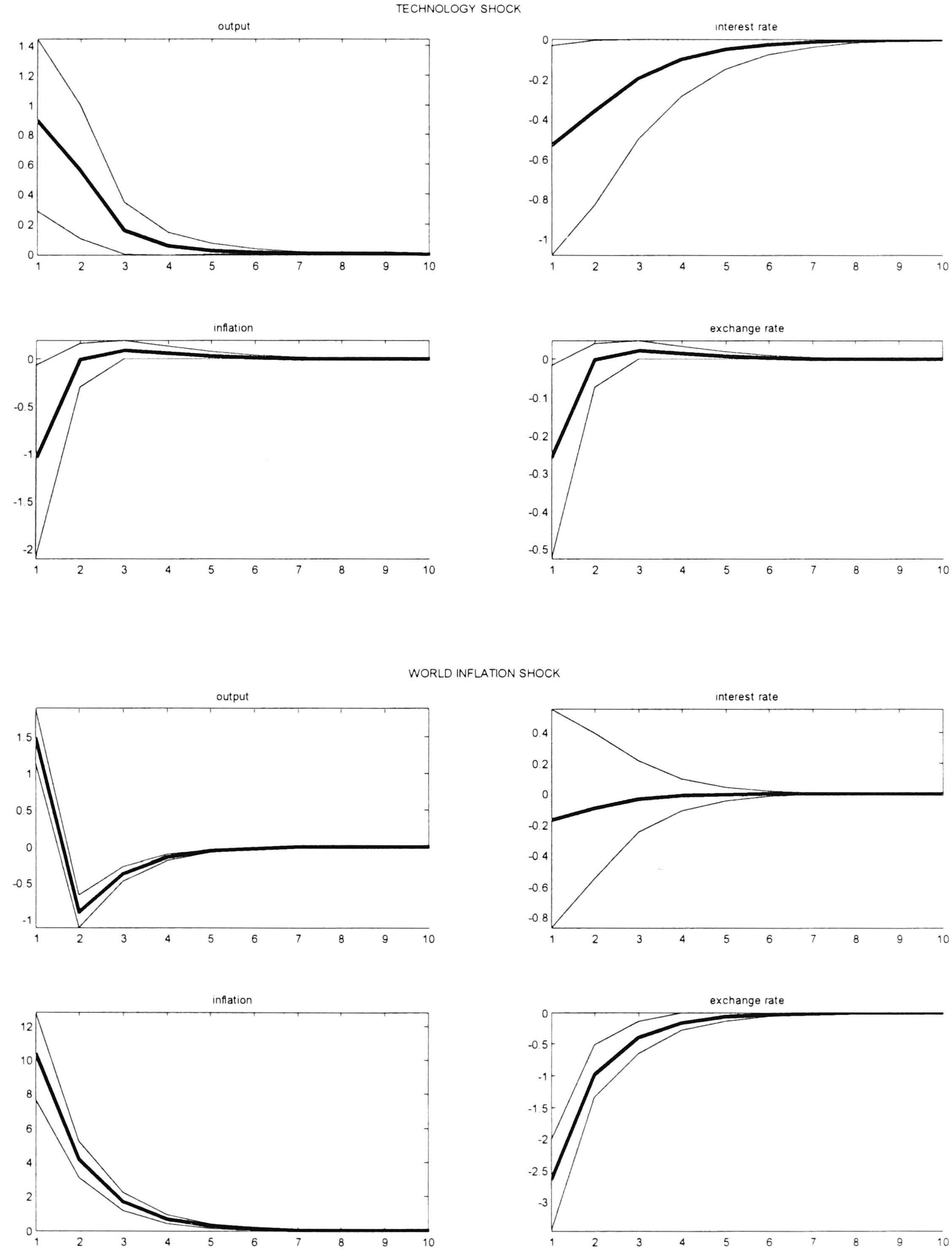

psi2 (Interval)
psi2 (m2)
psi2 (m3)
psi3 (Interval)
psi3 (m2)
psi3 (m3)
rho_R (Interval)
rho_R (m2)
rho_R (m3)

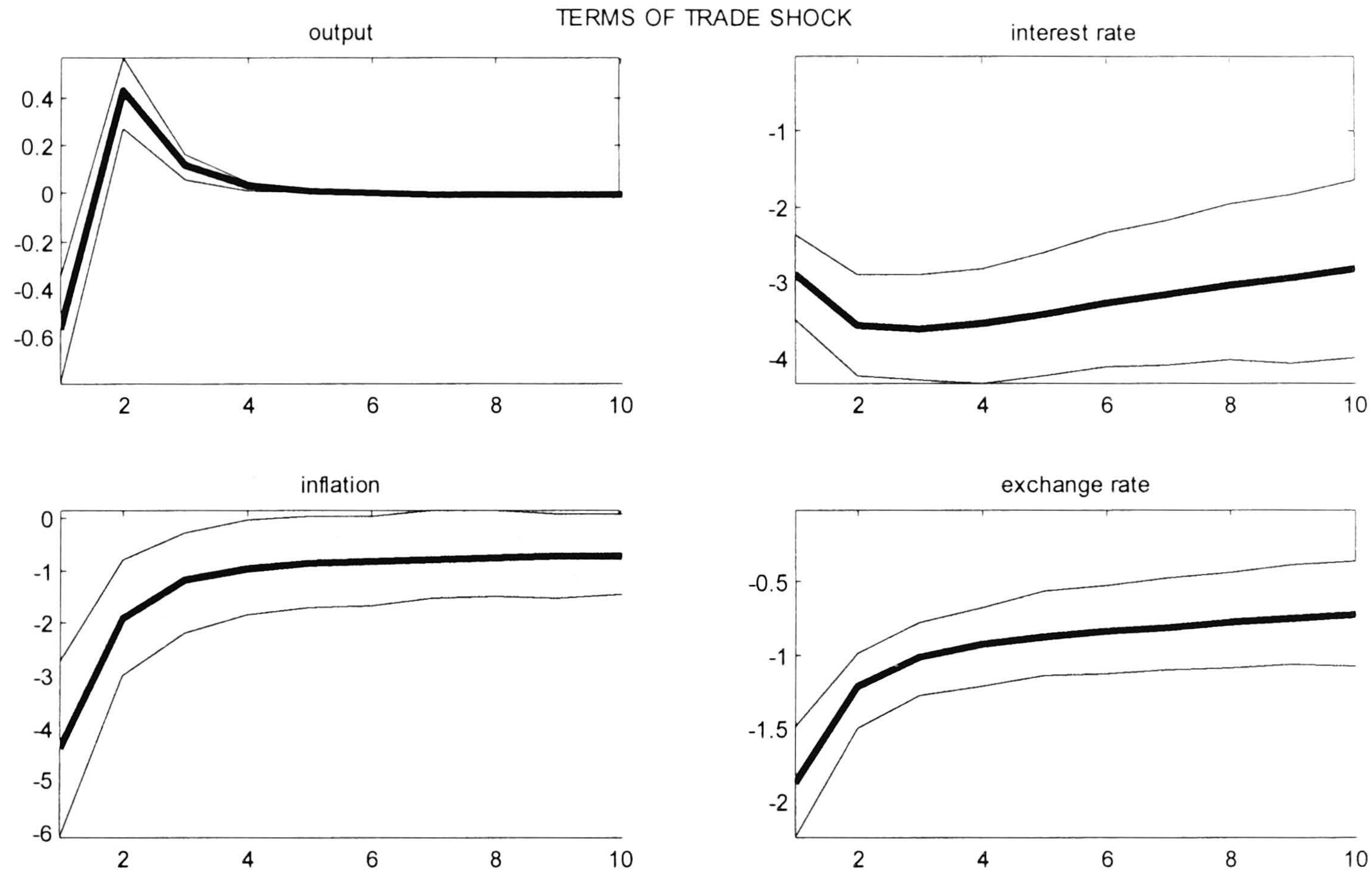

alpha (Interval)
alpha (m2)
alpha (m3)
rr (Interval)
rr (m2)
rr (m3)
k (Interval)
k (m2)
k (m3)

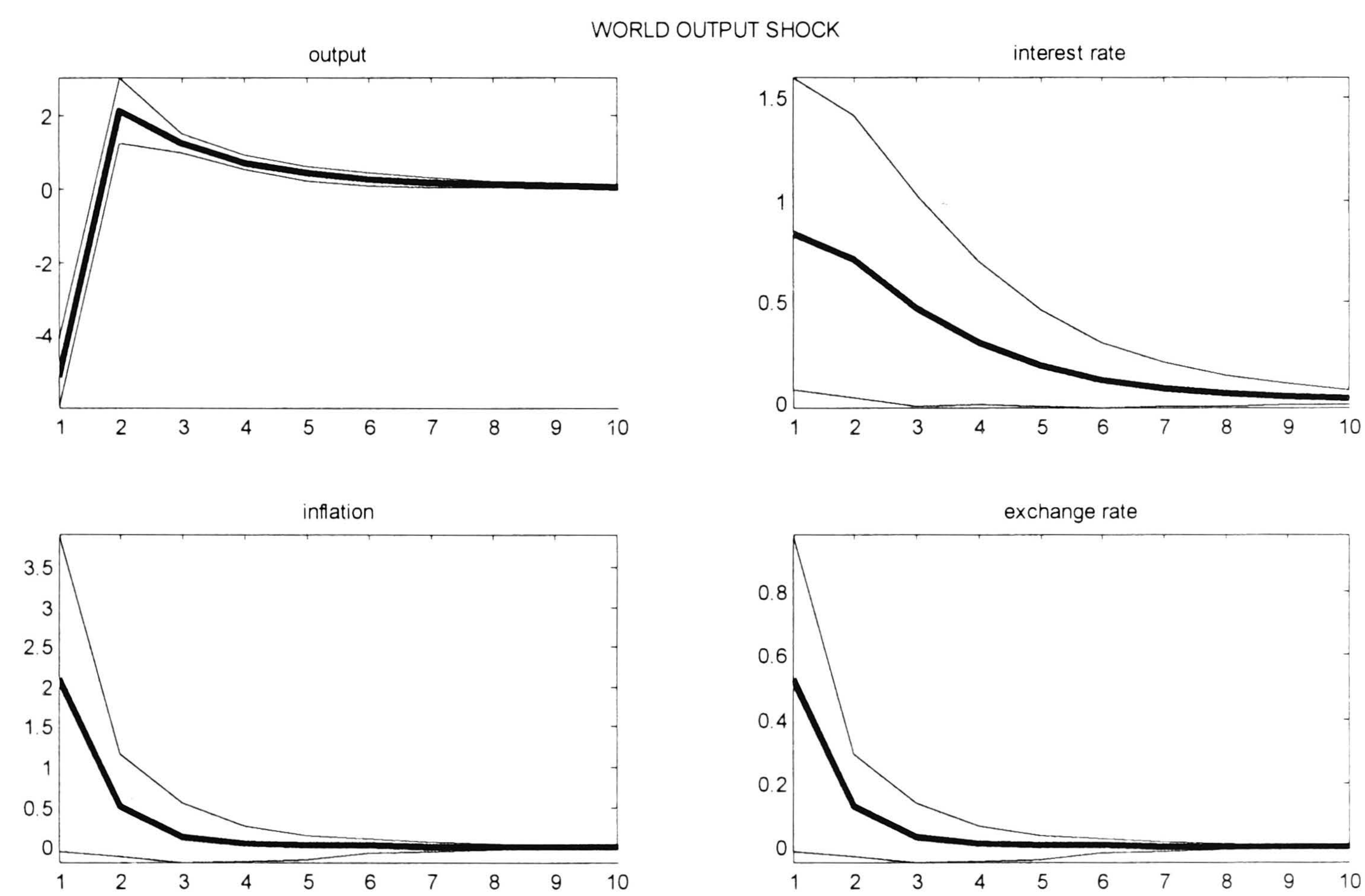


Source: Author's calculations

Figure 4.6a Impulse response functions for technology, world inflation, terms of trade and world output







Source: Author's calculations

Table 4.7a Variance Decomposition (%)

	Output	Inflation	Interest rate	Exchange rate
Policy shock	3.76	12.72	0.41	2.64
Terms of trade shock	1.51	26.68	99.29	85.07
Technology shock	3.15	0.72	0.08	0.15
World output shock	82.69	2.09	0.22	0.43
World inflation shock	8.9	57.79	0	11.7

Source: Author's calculations

Chapter 5 CONCLUSIONS

5.1 Summary of findings

The thesis consists of three essays. Chapter 2 attempts to measure the lagged effect of the monetary transmission mechanism on inflation and output in Mongolia using a sign restricted structural vector autoregression. We find the following results. First, the lag of the monetary transmission mechanism for Mongolia is about 4-12 months. Second, monetary policy shocks play a modest role in explaining output and inflation fluctuations. Third, in response to a monetary policy shock, the exchange rate immediately overshoots its long-run equilibrium.

Chapter 3 develops an empirical model for inflation in Mongolia using both Bayesian and classical approaches. In particular, we first estimate long-run markup and money demand relationships using the cointegration procedures, and then construct a single-equation error correction model of inflation with possible nonlinearity.

The main findings of Chapter 3 are summarized as follows. First, the main determinant of inflation is the markup, capturing impact from unit labor costs, petroleum prices, foreign inflation and the exchange rate. Second, money matters for inflation: excess narrow-money supply seems to cause inflation in the long run if the model uncertainty and nonlinearity are considered but adjustment to disequilibria is slow. Third, sustained increases in wages together with petroleum price shocks explain high and volatile inflation in recent years. We also find two inflationary regimes which are characterized by the degree of inflation persistence.

Chapter 4 estimates the reaction function of the Bank of Mongolia using a Bayesian approach. It addresses this issue by estimating the New Keynesian dynamic stochastic general equilibrium (DSGE) model of a small open economy. The main finding is summarized as follows. First, the monetary policy reaction function is forward-looking in terms of the inflation rate. Second, the central bank of Mongolia has implemented strong anti-inflationary and exchange rate stabilization policy. Third, there is evidence that the Bank of Mongolia does not respond to output. Finally, the DSGE-VAR analysis demonstrates that our estimated small open economy model can be used for forecasting output and inflation.

5.2 Policy implications

There are several implications of these findings for policy making in Mongolia. First, policy makers should not disregard the role of money in decision-making. The recent academic literature tends to neglect the exact role of money in inflation and output determination. For instance, money does not play any role in the New Keynesian model, which is the 'workhorse' for analysing monetary policy, and money is just a redundant variable. Even in practice, central banks tend to ignore information money contains for their policy making. However, money, especially narrow money, seems to have important information on price and output for Mongolia.

Second, the small models of structural vector autoregression, markup, money demand and New Keynesian open economy developed in the thesis could provide the starting point of medium or large scale models for Mongolia.

5.3 Limitations and suggestions for further studies

The empirical results of the essays in this thesis are subject to some limitations, especially with regard to the data availability and quality. We have several suggestions for further research. It would be useful to study the role of fiscal policy as a macroeconomic stabilization tool along with monetary policy. Furthermore, we need to study in detail the transmission of the foreign sector effects on the domestic economy since a small open economy such as that of Mongolia is vulnerable to external shocks.

REFERENCES

Albert, J.H. and Chib, S., 1993. Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. *Journal of Business and Economic Statistics*, 11, pp.1-15.

Amisano, G. and Giannini, C., 1997. *Topics in structural VAR econometrics*. Berlin: Springer-Verlag.

An, S. and Schorfheide, F., 2005. Bayesian analysis of DSGE model. *Econometric Reviews*, 26, pp.113–172.

Bauwens, L. and Giot., A., 1998. A Gibbs sampling approach to cointegration. *Computational Statistics*, 13, pp.339–368.

Bauwens, L. and Lubrano, M., 1996. Identification restrictions and posterior densities in cointegrated VAR systems. In T. B. Fomby, ed., *Advances in econometrics*, Volume 11, Part B, pp.3–28.

Benigno, G. and Benigno, P., 2003. Price stability in open economies. *Review of Economic Studies*, 70(4), pp.743–764.

Benigno, G. and Benigno, P., 2006. Designing targeting rules for international monetary policy coordination. *Journal of Monetary Economics*, 53(3), pp.473–506.

Bernanke, B.S., Gertler, M. and Watson, M.,1997. Systematic monetary policy and the effects of oil price shocks. *Brookings Papers on Economic Activity*, 1, pp.91–142.

Bernanke, B.S. and Mihov, I., 1998. Measuring monetary policy. *Quarterly Journal of Economics*, 113(3), pp.869–902.

Bjornland, H.C., 2009. Monetary policy and exchange rate overshooting: Dornbusch was right after all. *Journal of International Economics*, 79, pp.64–77.

Blanchard, O.J. and Quah, D.T., 1989. Dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79, pp.655–73.

Breunig, R. and Pagan, A., 2004. Do Markov-switching Models Capture Nonlinearities in the Data? Test Using Nonparametric Methods. *Mathematics and Computers in Simulation*, 64, pp. 401-407.

Brooks, S. and Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7, pp.434-455.

Bruggeman, A., Donati, P., and Warne A., 2003. *Is the demand for euro area M3 stable?* ECB working paper, p.255.

Calvo, G., 1983. Staggered prices in a utility maximising framework. *Journal of Monetary Economics*, 12(3), pp.383–398.

Campos, J., Ericsson, N.R. and Hendry, D.F., 2005. *General-to-specific Modeling: An Overview and Selected Bibliography*. International Finance Discussion Papers, 838, Board of Governors of the Federal Reserve System

Canova, F., 1991. The sources of financial crisis: Pre- and post-Fed evidence. *International Economic Review*, 32 (1991), pp.689–713.

Canova, F., 2007. *Methods for applied macroeconomic research*. Princeton, NJ: Princeton University Press.

Canova, F. and De Nicolo, G., 2002. Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics*, 49, pp.1131–1159.

Canova, F. and Menz, T., 2011. Does money matter in shaping domestic business cycles? An international evidence. *Journal of Money, Credit and Banking*, 43, pp.577–607.

Chib, S., 1995. Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association*, 90, pp.1313–1321.

Chib, S. and Greenberg, E., 1995. Understanding the Metropolis-Hastings algorithm. *American Statistician*, 49, pp.327–335.

Chistiano, L.J. and den Haan, W.J., 1996. Small-Sample Properties of GMM for Business-Cycle Analysis. *Journal of Business & Economic Statistics*, 14(3), pp.309-27.

Christiano, L.J., Eichenbaum, M. and Evans, C.L., 1999. Monetary policy shocks: what have we learned and to what end? *Handbook of Macroeconomics*, chapter 2, pp. 65-148 Elsevier.

Clarida, R., Gali, J. and Gertler, M., 1998. Monetary policy rules in practice: some international evidence. *European Economic Review*, 42, pp.1033–1067.

Clarida, R., Gali, J. and Gertler, M., 2001. Optimal monetary policy in open vs. closed economies: An integrated approach. *American Economic Review*, 91(2), pp.248–252.

Clarida, R., Gali, J. and Gertler, M., 2002. A simple framework for international monetary policy analysis. *Journal of Monetary Economics*, 49(5), pp.879–904.

Cooley, T. and LeRoy, S., 1985. A theoretical macroeconomics: A critique" *Journal of Monetary Economics*, 16, pp.283–330.

Corsetti, G. and Pesenti, P., 2005. International dimensions of monetary policy. *Journal of Monetary Economics*, 52(2), pp.281–305.

Cushman, D.O. and Zha, T., 1997. Identifying monetary policy in a small open economy under flexible exchange rates. *Journal of Monetary Economics*, 39(3), pp.433–448.

Davies, R.B., 1977. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrica*, 64, pp.247–54.

De Brouwer, G. and Ericsson, N.R., 1998. Modeling inflation in Australia. *Journal of Business and Economic Statistics*, 16-4, pp.433–49.

DeJong, D.N., Ingram, B.F. and Whiteman, C.H., 1993. Evaluating VARs with monetary business cycle priors. *Proceedings of the American Statistical Association*, Bayesian Statistical Science Section, pp.160–169.

Del Negro, M. and Schorfheide, F., 2004. Priors from equilibrium models. *International Economic Review*, 45, pp.663–673.

Devereux, M.B. and Engel, C., 2003. Monetary policy in the open economy revisited: Exchange rate flexibility and price setting behaviour. *Review of Economic Studies*, 70(4), pp.765–783.

Doan, T., 2010. *RATS 8.0 user's guide*, Evanston, IL: Estima.

Doan, T., Litterman R. and Sims C.A., 1984. Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews* 3, pp.1–100.

Dornbusch, R., 1976. Expectations and exchange rate dynamics. *Journal of Political Economy*, 84, 1 pp.161–1176.

Dungey, M. and Pagan, A., 2000. A structural VAR model of the Australian economy. *The Economic Record*, 76, pp. 321–342

Durevall, D. and Ndung'u, N.S., 1999. *A dynamic model of inflation for Kenya 1974-1996*. IMF working paper, 99/97

Faust, J., 1998. The robustness of identified VAR conclusions about money. *Carnegie-Rochester Series on Public Policy*, 49, pp.207–244.

Faust, J. and Leeper, E., 1997. When do long-run identifying restrictions give reliable results? *Journal of Business Economics and Statistics*, 15, pp.345–354.

Favara, G. and Giordani, P., 2009. Reconsidering the role of money for output, prices and interest rates. *Journal of Monetary Economics*, 56, pp.419–430.

Favero, C.A., 2001. *Applied macroeconometrics*. Oxford: Oxford University Press.

Fernandez, C., Ley, E. and Steel M.F.J, 2001. Benchmark priors for Bayesian model averaging. *Journal of Econometrics*, 100, pp.381–427.

Fernández-Villaverde, J., 2010. The econometrics of DSGE models. *SERIEs*, 1, pp.3-49.

Fernandez-Villaverde, J. and Rubio-Ramirez, J.F., 2004. Comparing dynamic equilibrium models to data: a Bayesian approach. *Journal of Econometrics*, 123, pp.153-187

Frühwirth-Schnatter, S., 2006. Finite mixture and Markov switching models. *Springer series in statistics*. New York/Berlin/Heidelberg: Springer

Fry, R and Pagan, A., 2011. Sign restrictions in structural vector autoregressions: A critical review. *Journal of Economic Literature*, 49, pp.938–960.

Gali, J. and Gertler M., 1999. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44, pp.195–222.

Gali, J. and Monacelli, T., 2005. Monetary policy and exchange rate volatility in small open economy. *Review of Economic Studies*, 72.

Gali, J., Gertler M. and Lopez-Salido, D., 2001. European inflation dynamics. *European Economic Review*, 45, pp.1237–1270.

Garcia, R., 1998. Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models. *International Economic Review* 39, pp.763-788

Geweke, J., 1996. Bayesian reduced rank regression in econometrics. *Journal of Econometrics*, 75, pp.121–146.

Geweke, J., 1999. Using simulation methods for bayesian econometric models: inference, development and communication. *Econometric Review*, 18(1), pp.1–126.

Gordon, D.B. and Leeper E.M., 1994. The dynamic impacts of monetary policy: An exercise in tentative identification. *Journal of Political Economy*, 102, pp.1228–1247.

Goyal, H., 1999. A development perspective on Mongolia. *Asian Survey*, 39, pp.633–655.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and business cycle. *Econometrica*, 57, pp.357–384.

Hamilton, J.D., 1994. *Time series analysis*. Princeton, NJ: Princeton University Press.

Hansen, B., 1992. Testing for parameter instability in linear models. *Journal of Policy Modelling*, 14, pp.517–533.

Hansen, H., and Johansen, S., 1999. Some test for parameter constancy in cointegrated VAR models. *Econometrics Journal*, 2, pp.306–333.

Hendry, D., 1995. On the Interactions of Unit Roots and Exogeneity, *Econometric Reviews*, 14-4, 383-419.

Hendry, D., 2000. Does money determine UK inflation over the long run? In: R. Backhouse and A. Salanti, eds. *Macroeconomics and the Real World*. Oxford: Oxford University Press, pp.85–114.

Ingram, B.F. and Whiteman, C.H., 1994. Supplanting the Minnesota prior – Forecasting macroeconomic time series using real business cycle model priors. *Journal of Monetary Economics*, 34, pp.497-510.

Johansen, S., 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, pp.231–244.

Johansen, S., 2002. Small sample correction of the test for cointegration rank in the vector autoregressive model. *Econometrica*, 70, pp.1929–1961.

Juselius, K., 1992. Domestic and foreign effects on prices in an open economy: The case of Denmark. *Journal of Policy Modeling*, 14, pp.401–428.

Kim, C.J., 1994. Dynamic linear models with Markov switching, *Journal of Econometrics*, 60, pp.1–22.

Kim, C.J. and Nelson, C.R., 1999. *State space model with regime switching: Classical and Gibbs-sampling approaches with applications*. Cambridge, MA: MIT Press.

Kleibergen F. and Paap, R., 2002. Priors, posteriors and bayes factors for a Bayesian analysis of cointegration. *Journal of econometrics*, 111, pp.223–249.

Kleibergen, F. and van Dijk, H.K., 1994. On the shape of the likelihood/posterior in cointegration models. *Econometric Theory*, 10, pp.514–551.

Kollmann, R., 2002. Monetary policy rules in the open economy: Effects on welfare and business cycles. *Journal of Monetary Economics*, 49(5), 989–1015.

Kuijs, L., 2002. *Monetary policy transmission mechanisms and inflation in the Slovak Republic*. IMF Working Paper 02/80.

Lane, P.R., 1999. The new open economy macroeconomics: Survey. *Journal of International Economics*, 54, pp.235–266.

Leeper, E.M., Sims, C.A. and Zha, T., 1996. What does monetary policy do? *Brookings Papers on Economic Activity*, pp.1–78.

Litterman, R.B., 1980. *A Bayesian procedure for forecasting with vector autoregression*. Department of Economics Working Paper, MIT.

Lubik, T.A. and Schorfheide, F., 2007. Do central banks respond to exchange rate movements? A structural investigation. *Journal of Monetary Economics*, 54, pp.1069–1087.

Lucas, R.E. Jr., 1976. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy,* 1, pp.19–46.

Mishkin, F., 1995. Symposium on the monetary transmission mechanism. *Journal of Economic Perspective*, 9:3, pp.3–10.

Monacelli, T., 2005. Monetary policy in a low pass-through environment. *Journal of Money, Credit and Banking*, 37(6), pp.1048–1066.

Mountford, A. and Uhlig, H., 2009. What are the effects of fiscal policy shocks?, *Journal of Applied Econometrics*, 24, pp.960–992.

Nyblom, J., 1989. Testing for the constancy of parameters over time. *Journal of American Statistical Association*, 84, pp.223–230.

Obstfeld, M. and Rogoff, K., 1995. Exchange rate dynamics redux. *Journal of Political Economy*, 103(3), pp.624–660.

Obstfeld, M. and Rogoff, K., 1996. *Foundations of International Macroeconomics.* Cambridge, MA:MIT Press.

Obstfeld, M. and Rogoff, K., 2002. Global implications of self-oriented national monetary policy rules. *Quarterly Journal of Economics*, 117(2), pp.503–535.

Pappa, E., 2004. Should the Fed and the ECB cooperate? Optimal monetary policy in a two-country world,. *Journal of Monetary Economics*, 51(4), pp.753–780.

Ploberger, W., Kramer, W. and Kontrus, K., 1989. New test for structural stability in the linear regression model. *Journal of Econometrics* 40, pp.307–318.

Rahbek, A., Hansen, E. and Dennis, J.G., 2002. *ARCH innovations and their impact on cointegration rank testing*. Working paper 22, Centre for Analytical Finance.

Rubio-Ramirez, J.F., Waggoner, D. and Zha, T., 2005 *Markov-switching structural vector autoregressions: theory and application*. Federal Reserve Bank of Atlanta Working paper, 2005-27

Sbordone, A., 2002. Prices and unit labor costs: A new test price stickiness. *Journal of Monetary Economics*, 54, pp.1311–1339.

Schorfheide, F., 2000. Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, 15(6), pp.645–670.

Sims, C.A.,1980. Macroeconomics and reality. *Econometrica*, 48, pp.1–48.

Sims, C.A., 1992. Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, 36, pp.975–1011.

Sims, C.A., 2011. Prize Lecture: Statistical Modeling of Monetary Policy and its Effects. Nobelprize.org.

Sims, C.A. and Uhlig, H.,1991. Understanding unit rooters: A helicopter tour. *Econometrica*, 59, pp.1591–1599.

Sims, C.A. and Zha, T., 1998. Bayesian methods for dynamic multivariate models. *International Economic Review*, 39(4), pp.949–968.

Sims, C.A. and Zha, T., 1999. Error bands for impulse response. *Econometrica*, 67(5), pp.1113–1156.

Slok, T., 2002. Money demand in Mongolia: A panel data analysis. *IMF Staff Papers*, 49, pp.128–135.

Smets, F. and Wouters, R., 2003. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5), pp.1123–1175.

Starr, M.A., 2005. Does money matter in the CIS? Effects of monetary policy on output and prices. *Journal of Comparative Economics*, 33(3), pp.441–461.

Strachan, R.W., 2003. Valid Bayesian estimation of the cointegrating error correction model. *Journal of Business & Economic Statistics*, 21, 185–195.

Strachan, R.W. and Inder, B., 2004. Bayesian analysis of the error correction model. *Journal of Econometrics*, 123, pp.307–325.

Sutherland, A., 2003. International monetary policy coordination and financial market integration. *Mimeo*, University of St. Andrews.

Uhlig, H., 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2), pp.381–419.

Villani, M., 2005. Bayesian reference analysis of cointegration. *Econometric Theory*, 21, pp.326–357.

Waggoner, D.F. and Zha, T., 2003. A Gibbs sampler for structural vector autoregressions. *Journal of Economic Dynamics and Control*, 28(2), pp.349–366.

Warne, A., 2006. *Bayesian inference in cointegrated VAR models: with applications to the demand for euro area*. ECB Working Paper, p.692.

Woodford, M., 2008. How important is money in the conduct of monetary policy? *Journal of Money, Credit and Banking*, 40–8, pp.1561–1598.

Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. New York: Wiley